1969

INTERMOLECULAR FORCES AND THE GROUP THEORY OF MANY-BODY SYSTEMS

P.E.S. WORMER

INTERMOLECULAR FORCES
AND THE
GROUP THEORY OF MANY-BODY SYSTEMS

Promotor:
Prof.Dr.Ir. A. van der Avoird

INTERMOLECULAR FORCES
AND THE
GROUP THEORY OF MANY-BODY SYSTEMS

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE WISKUNDE EN NATUURWETENSCHAPPEN AAN DE KATHOLIEKE UNIVERSITEIT TE NIJMEGEN, OP GEZAG VAN DE RECTOR MAGNIFICUS PROF. MR. F.J.F.M. DUYNSTEE VOLGENS BESLUIT VAN HET COLLEGE VAN DECANEN IN HET OPENBAAR TE VERDEDIGEN OP DONDERDAG 11 DECEMBER 1975 DES NAMIDDAGS TE 2.00 UUR PRECIES.

DOOR

PAUL EDUARD STEPHAN WORMER

GEBOREN TE AMSTERDAM

1975

DRUK: STICHTING STUDENTENPERS NIJMEGEN

Aan mijn moeder
en Jenny

CONTENTS

PREFACE

This thesis consists of two parts, which at first sight do not show much connection. However, the parts are related in that many of the group theoretical techniques, expounded in part I, have been applied in the calculations of part II. To account for the somewhat uncommon division of this dissertation into two parts, it is necessary to relate some of the history of the work.

My interest in the group theory of many-body systems was first aroused when I was working in Delft, under the supervision of Prof. L.L. van Reijen and in cooperation with Ir. J.J.M. Potters, on a spin-projected DODS (different orbitals for different spin) approach to the bonding in LiH. We worked within the AMO (alternant molecular orbital) scheme, which is a simplified version of the DODS method. (The results of this work are presented in ref. 15 of part I).

When later the work in Nijmegen on the calculation of intermolecular forces was started, we decided that it could be worthwhile to see what the AMO method would give for the London dispersion interaction between two He-atoms, because it is known that such an interaction is due to interatomic correlation and DODS methods are meant to give part of the correlation energy. So, the AMO programs were rewritten in FORTRAN (the original programs were in ALGOL), using the formula manipulation language FORMAC to derive and punch the really terrible formulas arising in a straightforward spin-projection (a Gelfand basis, discussed in sec. I.5.6, was used).

The AMO calculations gave remarkable results: using an atomic orbital basis consisting of s-orbitals only, a potential curve was obtained which closely resembled the experimental curves. This was most astonishing as it is known from London's theory (ch. II.3) that polarization functions (p-, d-, etc. orbitals) must be present on both the He-atoms to obtain a Van der Waals minimum. So we immediately knew the AMO results to be spurious. Furthermore, the AMO energies showed some unexpected regularities in their behaviour as a function of the variation parameters.

In order to explain these observations we had to look deeper into the transformation properties of spin-projected functions and to consider especially the behaviour of such functions under blocked orbital transformations, since the AMO method makes use of such a transformation.

(The results of this theoretical work are presented in two of the papers reprinted in this dissertation). It could now easily be proved that the He-He potential curve was due to a distance dependent intra-atomic correlation contribution to the AMO energy, thus showing that the AMO method was completely unsuited for this kind of work. The regularities in the energy as a function of the variation parameters could also be accounted for by the theory. In fact the AMO energy was proved to be a surprisingly simple function of these parameters, which also made clear that the formulas derived by FORMAC in a straightforward manner were unnecessarily complicated.

While doing this group theoretical work, I had to read parts of Weyl's and Boerner's books which I found an arduous job. But after having made this effort, it seemed that my experiences could be made useful by writing a survey that was self-contained and less condensed than the treatments in the books just mentioned. Since this review contained some material that could not be found explicitly in other places, we thought it worthwhile to include it in the first six chapters of this thesis.

Part II of this dissertation contains the results of the work on intermolecular forces which actually formed my main research interest during the past few years. For some introductory remarks on this part of the work the reader is referred to chapter II.1.

PART I:

GROUP THEORY OF MANY-BODY SYSTEMS

Because of the great complexity of the Schrödinger equation anyone performing molecular calculations is forced to introduce from the outset a number of simplifying assumptions, some of them rather severe, others not too strong. No quantum chemist will doubt the usefulness of the Born-Oppenheimer approximation or question the truncation of one-particle Hilbert space (orbital space). The Russell-Saunders coupling scheme, too, is generally accepted (except, of course, in cases where strong spin-orbit coupling occurs). Stated in mathematical terms these three widely applied approximations constitute the replacement of an exact Hamilton operator by a zeroth order N-electron model Hamiltonian, from which all energy contributions arising from spin are removed, in which all the nuclei are clamped, and which operates on a finite-dimensional Hilbert space that is a tensor product of N orbital spaces. This model is called the multi-configuration method, and includes most of the current computational models in quantum chemistry, such as valence bond, configuration interaction, and other methods.

It is after this point that consensus ceases to exist, a multitude of different approximations simplifying the model has been proposed, some of them so drastic that they virtually allow the Schrödinger equation to be solved by hand. However, after the advent of fast and large computers none of these further simplifications is in principle indispensable, at least not for simple molecules. Moreover, if computers and computing methods keep on developing at their present rate, the time is not far off when chemically more interesting systems will become tractable within the multi-configuration model. Hence it becomes a point of more than just academic interest to study the details of the model and to assess all its properties.

Fortunately we do not have to start from scratch; as so often before in physics, mathematicians have already paved the way. To clarify this we must note that, apart from possible spatial symmetry groups, two symmetry groups are always associated with the multi-configuration Hamiltonian, viz. the permutation group S_N and the special unitary group SU(2). Furthermore, the general linear group GL(n) enters the model in its rôle of dynamical group (a concept which is defined in chapter I.6). Of course, the representations of these three groups are not at all unrelated. As a matter of fact, one of the main purposes of the following exposition is the presentation of a coherent and unifying approach to the representation theory of

S_N and the Lie groups GL(n), U(n) and SU(n) together with their Lie algebras. These groups have indeed received much attention by mathematicians, and, after their bearing on quantum mechanics had been established, also by theoretical physicists.

The representation theory of S_N is linked with the names of Frobenius, Schur, Young, Littlewood and others. Weyl discussed the relation between S_N and the global representation theory of GL(n), whereas Wigner, Racah and other physicists concentrated mainly on the subgroup SU(2) of GL(2). Casimir, for instance, showed that the total spin angular momentum operator S^2 commutes with the elements of the Lie algebra of SU(2), thus tying the physical concept of spin with the mathematical concept of a Lie algebra, developed by Lie, Killing, Cartan, Weyl and others. During the last decade Biedenharn, Moshinksky, Louck and other workers have been generalizing the infinitesimal approach of SU(2) to SU(n) and GL(n), for arbitrary n. Clearly an enormous group theoretical machinery is in existence, ready for application in quantum chemistry.

However, the question may come to mind: do we really need all this group theory for handling the multi-configuration model? The traditional approach, which employs (spin-projected) Slater determinants is known to go a long way and has the advantage of not requiring any difficult mathematics. But if this traditional approach is pushed a little bit further, especially into the direction of atomic theory, one cannot avoid using shift operators, vector coupling coefficients etc. and so willy-nilly to employ Lie algebra theory, in particular the part that is related to SU(2). One may object that even if it is true that there is no way around SU(2), it does not necessarily mean that we must go further and introduce S_N and GL(n) also. There are some answers to that. In the first place it seems a waste not to explore an existing mathematical theory so ideally applying to our model. More important, however, is that we deepen our insight considerably by studying this theory; that is to say, the researcher who does the studying deepens his own personal understanding. For this to be useful for the whole of theoretical chemistry something more has to come out of it, though. What we would really like is a set of tools that cannot be furnished by the Slater determinant-plus-SU(2) approach alone. In atomic theory many instances of such results may be found, one only has to look into the books of Judd [1] or Wybourne [2] to find numerous examples. Especially the search, initiated by Racah, for complete sets of quantum

numbers has proved to be fruitful.

In the field of molecular quantum mechanics it is only very recently that people are beginning to realize that GL(n) and its unitary subgroup U(n) form a unifying framework from which arise such applications as for example:

- An algorithm for writing down the possible linearly independent anti-symmetric eigenstates of S^2.
- A closed formula for the total number of such states.
- Explicit relations for the transformation properties of such states under orbital mixings.
- Different sets of rules to evaluate matrix elements of the multi-configuration Hamiltonian.

Still, the full power of the Lie algebraic approach to GL(n) and U(n) remains to be explored with molecular applications in mind.

The permutation group has received a good deal more thought by quantum chemists over the years, with the attention focusing on the systematic derivation of the Hamilton matrix elements over states derived from non-orthogonal orbitals.

It is sometimes argued that group theory going beyond simple point groups is too abstract and complicated for use in chemistry. However, one must realize that every mathematical theory looks abstruse and complex before it has found general acceptance. To illustrate this point, I will relate some of the conceptual problems mathematicians experienced when letters carrying signs came into use for denoting unknowns in simple algebraic equations. For example, Hieronymus Cardanus (1501-1576), one of the leading mathematicians of his time, published in 1570 a "proof" of the fact that the rule $(-a)\cdot(-b) = +a\cdot b$ does not hold for letters. Mind you, this rule had been known for numbers since Diophantus' time (third century). Thirty eight years later another mathematician, Christophorus Clavius (1537-1612), still could not quite grasp the use of letters, as is witnessed by the following citation from his famous textbook (translated from the Latin text quoted by Tropfke [3]):

"It seems that one must forgo the reason why the unknowns and their signs multiply as they do: this must be attributed to the shortcomings of the human mind that cannot comprehend in which manner it can be possibly true."

One of the main lessons that can probably be drawn from the history of science, is not to be deterred by an apparently esoteric theory.

The first four chapters of the subsequent review contain a rather thorough discussion concerning the tensor representations of GL(n), U(n) and SU(n) and the relation with the corresponding theory for the respective Lie algebras. Also the connection with the representation theory of S_N is discussed in detail. Much of this material can be found in the books of Weyl [3, 4] and Boerner [5]. These books being written in a highly condensed style, the main objective of the present work is to give a self-contained account that is readable by the average quantum chemist, without demanding too much effort. To that end two introductory chapters, one on associative algebras and one on linear Lie groups, precede the chapters I.4 and I.5, which together form the backbone of the present review.

The relation between electron spin, SU(2), rotational and permutational symmetry is gradually worked out in a number of examples. In chapter I.6 the bearing of the foregoing chapters on the quantum mechanics of N-particle systems is discussed.

For reasons of space many of the proofs are omitted, but always with a reference (including page or theorem number) to a source where the proof can be found. The proofs that are given explicitly satisfy one or both of the two criteria:

- The theorem is of such importance that an explicit proof, clarifying the content of the theorem, is justified.
- No proof could be found in the literature.

When writing a review such as this it is unavoidable that one comes across numerous blank spots that must be filled out in the course of writing. For instance the following points do not come directly from other sources:

- A proof that every linear operator on tensor space, commuting with all permutations, can be written as a polynomial in the generators of GL(n) (sec. I.4.5).
- A proof that the (tensor) Casimir operators of GL(n) can be expressed in terms of the class sum operators of S_N (sec. I.4.7).
- An explicit and simple formula for the matrix representations of S_N carried by Young units (sec. I.5.1).
- An algebraic proof of the fact that Young orthogonal units and Yamanouchi units are identical (sec. I.5.3).
- A proof of the completeness and linear independence of the set of standard Weyl-Rumer tensors (sec. I.5.5).

- A proof of the completeness and linear independence of the set of standard Gelfand tensors (sec. I.5.6).
- An explicit formula for the matrix representations of GL(n) carried by Gelfand or Weyl-Rumer tensors (sec. I.5.7).
- The discussion of N-particle symmetry (sec. I.6.2).
- The realization that GL(n) is a dynamical group of the Hamiltonian (sec. I.6.3).
- The antisymmetrizer expressed in terms of Young units (sec. I.6.4).
- The explicit relation between the conventional approach via spin-bonded functions and the spin-free approach using Weyl-Rumer tensors (sec. I.6.5).

Of course many of these points are already implicitly contained in the work of others, and not having made a complete literature survey I may have easily missed earlier explicit proofs of the same facts. In this connection a series of stimulating lectures by professor Matsen,given at this university in October 1974, must especially be mentioned. That was when my attention was first drawn to several of the points listed above.

The review is finally followed by a few papers published earlier. Each of these carries its own introduction; the reader is referred to those for the purpose and scope of these articles.

CHAPTER I.2. A REVIEW OF SOME ALGEBRA

This chapter summarizes some abstract algebra which is required in the later sections, and which is probably not general knowledge among chemists. Not reviewed are the following better known subjects, also having much bearing on the first half of this thesis:

- finite groups and their irreducible representations; for instance expounded in ref. 4, chapters I.1 - I.5 and II.1 - II.5;
- linear operators (and their matrices) on finite dimensional vector spaces (see e.g. ref. 5, chapters I, II, V and VI).

Most of the theory of this chapter can be found scattered and interwoven with other material in the books of Weyl [6, 7] and Boerner [8]. References 9, 10 and 11 should be consulted especially for the properties of semisimple associative algebras, reviewed briefly in sec. I.2.2.

2.1. Definition of a few different algebras

Definition.

A set A is an associative algebra over the field of complex numbers C if it is a vector space over C and has an internal binary operation (multiplication) satisfying the properties:

A1. $a(b + c) = ab + ac$; $(b + c)a = ba + ca$;

A2. $a(bc) = (ab)c$ (the associativity property);

A3. $(\lambda a)b = \lambda(ab) = a(\lambda b)$;

for all $a, b, c \in A$ and $\lambda \in C$.

Definition.

A set Λ is a Lie algebra over the field of real numbers R if it is a vector space over R and has an internal binary operation (commutator bracketing) satisfying the properties:

L1. $[a, b + c] = [a, b] + [a, c]$; $[b + c, a] = [b, a] + [c, a]$;

L2. $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$;

L3. $[a, a] = 0$;

L4. $[\lambda a, b] = \lambda[a, b] = [a, \lambda b]$;

for all $0, a, b, c \in \Lambda$ and $\lambda \in R$.

When the field R is extended to C, Λ becomes a complex Lie algebra: the complexification of Λ.

Theorem.

$[a, b] = -[b, a]$.

Proof. Substitute $a = a + b$ in L3 and use L1 and L3.

Notes.

1. Although it is not strictly necessary, we will assume an algebra (Lie or associative) to be finite-dimensional in the vector space sense. Further we assume associative algebras to have a unit element.
2. In the sequel we will almost exclusively be dealing with algebras of linear operators. Linear operators on a vector space can be multiplied by a scalar, added and multiplied, in all cases yielding again a linear operator. (The multiplication of linear operators is associative). Thus, if a set of linear operators is closed under these three algebraic operations the set is an associative algebra.
3. The commutator bracket:

$$[A, B] = AB - BA$$

of two linear operators A and B is easily shown to satisfy the requirements L1 - L4. Thus, if a set of operators is closed under multiplication by a scalar, addition and commutator bracketing, the set is a Lie algebra.

Definition.

Let $\{a_i\}$ be a basis of the associative algebra A. Since $a_i a_j \in A$ we can write:

$$a_i a_j = \sum_k \gamma_{ijk} a_k, \qquad \gamma_{ijk} \in C.$$

The constants γ_{ijk} are the structure constants of A. Analogously one defines the structure constants of Lie algebras.

Example.

Define the following three matrices, representing the three components of the angular momentum operator $\vec{L}$:

$$\underline{L}_x = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \underline{L}_y = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \underline{L}_z = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The well-known commutation relations of angular momentum operators yield:

$$[\underline{L}_i, \underline{L}_j] = i \sum_{k=1}^{3} \epsilon_{ijk} \underline{L}_k, \qquad i, j, k = x, y, z = 1, 2, 3.$$

The constants ε_{ijk} are the antisymmetric Levi-Civita symbols defined by.

$\varepsilon_{ijk} = +1$ if $i < j < k$ and any even permutation of i, j and k.

$= -1$ for an odd permutation of i, j and k.

$= 0$ if two or more indices are equal.

The 3-dimensional linear space generated by $\underline{L}_x$, $\underline{L}_y$, $\underline{L}_z$ is a (matrix) Lie algebra with structure constants $i\varepsilon_{ijk}$. Later we will show this algebra to be the Lie algebra of SO(3) (sec. I.3.1).

Theorem.

The set of all linear operators on an n-dimensional vector space V_n is an n^2-dimensional associative algebra, denoted by AL(n).

Proof. Ref. 5, theorems 2.6 and 2.8.

Linear operators on an n-dimensional space are faithfully represented by n × n matrices. The set of all n × n matrices AL(n,C) is therefore an n^2-dimensional associative algebra too.

Definition.

Let S be a subset of an associative algebra A. Close S multiplicatively. That is, augment S with all possible different products of elements in S, where the product is the associative product of A. This yields a new subset S' of A. the multiplicative closure of S (a semigroup). Addition and multiplication by scalars are defined on S', since it is a subset of A. Close S' linearly with respect to these vector space operations. That is, form the vector space $\mathfrak{G}(S)$ generated by the maximal number of linearly independent elements in S'. The algebra $\mathfrak{G}(S)$, the smallest subalgebra of A containing S, is the enveloping algebra of S (in A).

Example.

Let $\Lambda \subset AL(n)$ be the set of all skew-Hermitean operators. We will show in sec. I.3.2 that Λ is a Lie algebra over R, and hence is by definition linearly closed. Extending the field R to C, Λ ceases to be closed, for iA, $A \in \Lambda$, is Hermitean. It is easily shown that $i\Lambda$ is the set of all Hermitean operators. Since it is quite easy to prove that every linear operator can be written uniquely as the sum of a skew-Hermitean and a Hermitean operator, we find:

$$AL(n) = \Lambda \oplus i\Lambda$$

and so AL(n) is the enveloping algebra (over C) of Λ. Because, as we will see later (sec. I.3.2), Λ is the Lie algebra of U(n), and because AL(n) is the Lie algebra of GL(n) (sec. I.3.2) we have found the interesting result that the Lie algebra of GL(n) is the complexification of the Lie algebra of U(n). We will use this fact later, when discussing Weyl's "Unitarian Trick" (sec. I.3.2).

Definition.

Let G be a finite abstract group of order ^{o}G. Since by definition G is already closed under multiplication, its linear closure, consisting of the formal sums:

$$\sum_{g \in G} c(g)g, \qquad g \in G,\ c(g) \in C,$$

is an algebra of dimension ^{o}G. This algebra is the group algebra of G, denoted by CG.

Note.

In this definition the adjective 'formal' is essential (see e.g. ref. 12, p. 4 for a rigorous definition of the formal sum). If it happens that addition and multiplication by scalars have already been defined on G, then the formal sum must be clearly distinguished from these operations. For example, on the group O(3), consisting of all 3 × 3 real orthogonal matrices, addition and multiplication by a scalar are defined as is usual for matrices. The enveloping algebra of O(3) in AL(3) is finite-dimensional. However, the group algebra of O(3) is of infinite dimension, since O(3) is of infinite order. (In fact this group algebra is a pre-Hilbert space; see ref. 13, p. 408).

Definitions.

1. Let ϕ be a mapping of a group G into a group G', then ϕ is a group homomorphism if it satisfies:
$$\phi(gg') = \phi(g)\phi(g'), \qquad g, g' \in G.$$
2. Let ϕ be a mapping of an algebra A into an algebra A', then ϕ is an algebra homomorphism if it satisfies:
$$\phi(aa') = \phi(a)\phi(a')$$
$$\phi(\lambda a + \mu a') = \lambda\phi(a) + \mu\phi(a'), \qquad a, a' \in A;\ \lambda, \mu \in C.$$
If in these two cases ϕ is a 1-1 mapping and onto, we speak of a group or an algebra isomorphism respectively.

Notes.

1. From here on only enveloping algebras of sets of linear operators on n-dimensional spaces will be considered. It is understood that these algebras are always imbedded in AL(n), without mentioning this explicitly.
2. The enveloping algebra of an n-dimensional irrep D of a finite group G is always equal to the full matrix algebra AL(n). This follows from Burnside's theorem [8, p. 65], which states that D(G) containes exactly n^2 linearly independent operators.

We will find several examples in the sequel where the irreducibility of a group representation is proved by first ascertaining the same fact for its enveloping algebra. The validity of this procedure derives from the following theorem, holding for general groups.

Theorem.

The enveloping algebra $\mathfrak{E}(G)$ of an operator group G is irreducible if and only if G is irreducible.

Proof. Apply the argument of ref. 9, p. 45.

Commutator algebra.

If S is an arbitrary set of linear operators on a vector space V, it is easy to show that the set:

$$S^C = \{s^C \mid s^C s = s s^C, \forall s \in S\}$$

consisting of all operators on V that commute with the elements of S is an algebra. The algebra S^C is called the intertwining, commutator or commuting algebra of S.

2.2 Semi-simple associative algebras

Almost all associative algebras encountered in quantum mechanics have completely reducible representations. Just as is the case with groups, this property of the representation is solely determined by the structure of the algebra itself. Unfortunately it requires too many pages to present something remotely resembling a self-contained account of the structure of algebras, and so we have to make do with only a sketch of the theory.

Definitions.

1. A linear subspace L of an associative algebra A is a left ideal if it is stable under left multiplication with the elements of A. So, L has the property:

$$aL \subset L, \qquad \forall a \in A.$$

2. If L does not contain any left ideals except itself and (0), it is a minimal left ideal.
3. In an analogous way one defines a (minimal) right ideal.
4. An element $e \in A$ is an idempotent if e has the property: $e^2 = e$.
5. Two elements $a', a'' \in A$ are orthogonal if $a'a'' = 0$.
6. Two subsets S' and S" of A are orthogonal if $S'S'' = (0)$; that is: $s's'' = 0$, $\forall s' \in S'$ and $\forall s'' \in S''$.
7. An idempotent e is primitive if no two non-zero and orthogonal idempotents e' and e" exist, such that $e = e' + e''$.

The set Aa, for some arbitrary fixed $a \in A$, is a left ideal generated by a.

Theorem.

Let L be the left ideal generated by the idempotent e. L is minimal if, and only if, e is primitive.

Definitions.

1. A set $A' \subset A$ which is simultaneously a left and a right ideal in A is a two sided ideal (other names. ideal, invariant subalgebra).
2. A finite dimensional algebra that contains no ideals except itself and (0) is a simple algebra.

Theorem.

(i) The unit e of a simple algebra A can be decomposed thus

$$e = e_1 + e_2 + \ldots + e_f,$$

where the e_i are primitive orthogonal idempotents.

(ii) A simple algebra A is a direct sum of orthogonal minimal left ideals generated by the e_i:

$$A = L_1 \oplus L_2 \oplus \ldots \oplus L_f,$$

where the left ideals are all of the dimension f; so, A is of dimension f^2.

Theorem: (Wedderburn).

(i) The minimal left ideal L_i occurring in the decomposition of a simple algebra A has a basis: $\{e_{ji} \mid j = 1, \ldots, f\}$, $i = 1, \ldots, f$. This basis has the property:

$$e_{ij}\, e_{k\ell} = \delta_{jk}\, e_{i\ell}, \qquad i, j, k, \ell = 1, \ldots, f.$$

(ii) If $a \in A$ is expressed with respect to this basis as follows:

$$a = \sum_{k,\ell} D(a)_{k\ell}\, e_{k\ell}$$

then:

$$ae_{ij} = \sum_{k} D(a)_{ki}\, e_{kj}, \qquad a \in A,\ D(a)_{ki} \in C.$$

(iii) The set $\{\underline{D}(a) \mid a \in A\}$ comprises all $f \times f$ matrices and is therefore the full matrix algebra AL(f,C), which hence is a faithful representation of A.

Note.

The matrix representation $\underline{D}(a)$ of $a \in A$ does not depend on the choice of the minimal left ideal that carries the representation.

Example.

The full matrix algebra AL(n,C) is the best known example of a simple associative algebra; it is of the dimension n^2. The subset L_i, consisting of matrices with zeros everywhere except in column i, is an n-dimensional minimal left ideal with basis $\{\underline{E}^{ji} \mid j = 1, \ldots, n\}$, where $\underline{E}^{ji}$ is a matrix with all elements equal to zero, except the (j,i)-element,which is equal to one. The decomposition of AL(n,C) into minimal left ideals is obvious, the corresponding decomposition of the unit is:

$$\underline{E} = \sum_{i=1}^{n} \underline{E}^{ii}.$$

One can proceed analogously with the minimal right ideals, which are sets of matrices with one row filled and zeros elsewhere.

Definitions.

1. An algebra A is semi-simple if it is a direct sum of simple algebras:

$$A = \sum_{\lambda} \oplus\, A^{(\lambda)},$$

with:

$$A^{(\lambda)} A^{(\mu)} = (0).$$

The second condition implies that the representations of A afforded by the different simple algebras in this decomposition are non-equivalent.

2. The simple algebras $A^{(\lambda)}$ are called the simple components of A.

Theorem.

The unit e of a semi-simple algebra A can be written as:

$$e = \sum_{\lambda} e^{(\lambda)}$$

where $e^{(\lambda)}$ is the unit of the simple component $A^{(\lambda)}$ of A.

The following theorem provides a criterion enabling us to decide whether an algebra is semi-simple.

Theorem.

An algebra A is semi-simple if and only if each of its ideals is direct summand. That is, if A' is an ideal in A, there exists always an ideal A" in A, such that $A = A' \oplus A"$.

Examples.

1. Let A be a decomposable operator algebra on a vector space V. This means that each A-invariant subspace of V has an A-invariant complement in V. So, if $U \subset V$ is A-invariant, there exists an A-invariant space $W \subset V$ such that $V = U \oplus W$. An element $\alpha \in A$ decomposes accordingly as: $\alpha = \alpha_U \oplus \alpha_W$, where α_U and α_W are the restriction of α to U and W, respectively. The sets $A_U = \{\alpha_U \oplus 0\}$ and $A_W = \{0 \oplus \alpha_W\}$ are ideals in A with the property:

$$A = A_U \oplus A_W.$$

So, a decomposable algebra is semi-simple.

2. The most important example of a semi-simple algebra is the group algebra CG of a finite group G. Recall [14] that CG has a basis of elements ("Wigner operators") of the form:

$$e_{ij}^{(\lambda)} = \sum_{g \in G} D^{(\lambda)}(g^{-1})_{ji}\, g,$$

where D is an irreducible matrix representation of G.
The set $\{e_{ij}^{(\lambda)} \mid i, j = 1, \ldots, f_{(\lambda)};\ \lambda \text{ fixed}\}$ spans the simple component $A^{(\lambda)}$ of CG. The element

$$e^{(\lambda)} = \sum_{i=1}^{f^{(\lambda)}} e_{ii}^{(\lambda)} = \sum_{g \in G} \chi^{(\lambda)}(g^{-1})\, g,$$

where $\chi^{(\lambda)}(g^{-1})$ is the trace of $\underline{D}^{(\lambda)}(g^{-1})$, is the unit of $A^{(\lambda)}$. The unit $e^{(\lambda)}$ is often called a "character projector".

3. Later in this chapter it will be shown that the commutator algebra of a semi-simple operator algebra is semi-simple too.

The following two theorems concern the representations of semi-simple algebras. The first is the analogue of Maschke's theorem [4, p. 95], and the second is a generalization of a well-known theorem holding for the regular representation of finite groups [4, p. 108].

Theorem.
Every linear representation of a semi-simple algebra is decomposable.

Theorem.
Every linear irreducible representation (irrep) of a semi-simple algebra is equivalent to one of the irreps carried by the algebra itself.

2 3 Representations of an algebra and its commutator algebra

In this section we will consider how V decomposes under two sets of commuting operators, and see under which conditions the decompositions involved are complete.

Definition.
The Kronecker product $\underline{R} \otimes \underline{S}$ of two arbitrary matrices $\underline{R}$ and $\underline{S}$ is defined as the matrix:

$$\underline{R} \otimes \underline{S} = \begin{pmatrix} R_{11}\underline{S} & R_{12}\underline{S} & \cdots & R_{1m}\underline{S} \\ R_{21}\underline{S} & \vdots & & \vdots \\ \vdots & & & \vdots \\ R_{n1}\underline{S} & \cdots & & R_{nm}\underline{S} \end{pmatrix}$$

So $\underline{R} \otimes \underline{S}$ is a matrix with the matrix $R_{ij}\underline{S}$ at position (i,j). Another way of looking at $\underline{R} \otimes \underline{S}$ is to regard it as a matrix with its elements (scalars) labelled by four indices

$$(\underline{R} \otimes \underline{S})_{ij;k\ell} = R_{ik} S_{j\ell}.$$

The following multiplication rule for Kronecker products can be easily proved:

$$(\underline{R} \otimes \underline{S})(\underline{R}' \otimes \underline{S}') = (\underline{RR}') \otimes (\underline{SS}').$$

Note.

Kronecker products can also be defined for more than two factors, postulating associativity.

Now we state an important theorem, which will be one of the cornerstones of the theory expounded in I.4.

Theorem.

Let A_n be a semi-simple algebra of operators on the vector space V_n. Let V_n be decomposed in A_n-irreducible subspaces:

$$V_n = \sum_{\lambda=1}^{p} \sum_{j=1}^{f^{(\lambda)}} \oplus U_j^{(\lambda)} \tag{1}$$

Let the spaces $U_j^{(\lambda)}$, $j = 1, \ldots, f^{(\lambda)}$ carry identical irreducible matrix representations of dimension $n^{(\lambda)}$.

If A_n^c is the commutator algebra of A_n, then we will prove that a basis of V_n, adapted to the decomposition (1), is simultaneously adapted to the decomposition:

$$V_n = \sum_{\lambda=1}^{p} \sum_{i=1}^{n^{(\lambda)}} \oplus W_i^{(\lambda)} \tag{2}$$

The spaces $W_i^{(\lambda)}$, $i = 1, \ldots, n^{(\lambda)}$ carry identical irreducible $f^{(\lambda)}$-dimensional matrix representations of A_n^c.

Note that the frequency factor $f^{(\lambda)}$ and the dimension $n^{(\lambda)}$ have changed rôles on going from decomposition (1) to (2).

Proof. Let

$$\{u_{k\ell}^{(\lambda)} \mid \lambda = 1, \ldots, p;\ k = 1, \ldots, f^{(\lambda)};\ \ell = 1, \ldots, n^{(\lambda)}\}$$

be a basis of V_n corresponding to decomposition (1). The matrix $\underline{A}$ of $A \in A_n$ is defined by:

$$Au_{k\ell}^{(\lambda)} = \sum_{\mu=1}^{p} \sum_{i=1}^{f^{(\lambda)}} \sum_{j=1}^{n^{(\lambda)}} A_{ij;k\ell}^{(\mu;\lambda)} u_{ij}^{(\mu)}.$$

Since $U_k^{(\lambda)}$ is stable under $A \in A_n$, the matrix of A takes the form:

$$A^{(\mu;\lambda)}_{ij;k\ell} = \delta^{\mu,\lambda}\, \delta_{ik}\, D^{(\lambda)}(A)_{j\ell}.$$

The matrix $\underline{D}^{(\lambda)}(A)$ represents A on $U_k^{(\lambda)}$. Clearly

$$\delta_{ik}\, \underline{D}^{(\lambda)}(A)$$

is the (i,k)-element of

$$\underline{E}_{f^{(\lambda)}} \otimes \underline{D}^{(\lambda)}(A),$$

where $\underline{E}_{f^{(\lambda)}}$ is the $f^{(\lambda)} \times f^{(\lambda)}$ unit matrix. So, the matrix $\underline{A}$ of $A \in A_n$ can be written as:

$$\underline{A} = \sum_{\lambda=1}^{p} \oplus\, \underline{E}_{f^{(\lambda)}} \otimes \underline{D}^{(\lambda)}(A),$$

or alternatively:

$$\underline{A} = \underline{E}_p \otimes \underline{E}_{f^{(\lambda)}} \otimes \underline{D}^{(\lambda)}(A),$$

where $\underline{E}_p$ is the $p \times p$ unit matrix.

We will now show that the matrix $\underline{B}$ of $B \in A_n^c$ has a very similar form. From $AB - BA = 0$ follows $\underline{AB} - \underline{BA} = \underline{0}$, or:

$$\sum_{\nu=1}^{p} \sum_{q=1}^{n^{(\lambda)}} \sum_{r=1}^{f^{(\lambda)}} \left[A^{(\mu;\nu)}_{ij;rq}\, B^{(\nu;\lambda)}_{rq;k\ell} - B^{(\mu;\nu)}_{ij;rq}\, A^{(\nu;\lambda)}_{rq;k\ell} \right] = 0$$

and substituting the special form of $\underline{A}$ it follows that:

$$\sum_{q=1}^{n^{(\lambda)}} \left[D^{(\mu)}(A)_{jq}\, B^{(\mu;\lambda)}_{iq;k\ell} - B^{(\mu;\lambda)}_{ij;kq}\, D^{(\lambda)}(A)_{q\ell} \right] = 0,$$

which reads in matrix notation:

$$\underline{D}^{(\mu)}(A)\, \underline{B}^{(\mu;\lambda)}_{i;k} - \underline{B}^{(\mu;\lambda)}_{i;k}\, \underline{D}^{(\lambda)}(A) = \underline{0},$$

where $\underline{B}^{(\mu;\lambda)}_{i;k}$ is an $n^{(\lambda)} \times n^{(\lambda)}$ matrix with the element $B^{(\mu;\lambda)}_{iq;k\ell}$ at the (q,ℓ) position.

Now by Schur's lemma [4, p. 98] we get:

$$\underline{B}^{(\mu;\lambda)}_{i;k} = c\underline{E}_{n^{(\lambda)}} \qquad \text{if } \lambda = \mu$$

$$= \underline{0} \qquad \text{if } \lambda \neq \mu.$$

The Schur constant c depends on B, λ, i and k, and so we write $D^{(\lambda)}(B)_{ik}$ rather than c, thus making clear that these constants form an $f^{(\lambda)} \times f^{(\lambda)}$ matrix, when keeping B and λ fixed.

The matrix $\underline{B}$ of B can finally be written as:

$$\underline{B} = \sum_{\lambda=1}^{p} \oplus \underline{D}^{(\lambda)}(B) \otimes \underline{E}_{n(\lambda)}$$

or alternatively:

$$\underline{B} = \underline{E}_p \otimes \underline{D}^{(\lambda)}(B) \otimes \underline{E}_{n(\lambda)}.$$

Up to this point we never used the fact that A_n^c is an algebra, and indeed the result just found holds for an arbitrary set of operators commuting with A_n. To prove, as the next step, the irreducibility of $\{\underline{D}^{(\lambda)}(B)\}$ we do need the fact that A_n^c contains <u>all</u> operators commuting with A_n, which makes A_n^c automatically an algebra.

Every operator on V_n with a matrix of the form:

$$\underline{E}_p \otimes \underline{C} \otimes \underline{E}_{n(\lambda)}, \qquad \text{arbitrary } \underline{C} \in AL(f^{(\lambda)}, C),$$

which is defined with respect to the basis $\{u_{k\ell}^{(\lambda)}\}$, belongs to A_n^c. This follows from the commutation of such a matrix with the matrices of A_n:

$$(\underline{E}_p \otimes \underline{C} \otimes \underline{E}_{n(\lambda)})(\underline{E}_p \otimes \underline{E}_{f(\lambda)} \otimes \underline{D}^{(\lambda)}(A)) = (\underline{E}_p \otimes \underline{E}_{f(\lambda)} \otimes \underline{D}^{(\lambda)}(A))(\underline{E}_p \otimes \underline{C} \otimes \underline{E}_{n(\lambda)}).$$

Hence the matrix $\underline{E}_p \otimes \underline{C} \otimes \underline{E}_{n(\lambda)}$ is the matrix of an element in A_n^c.

Since this holds for <u>all possible</u> $f^{(\lambda)} \times f^{(\lambda)}$ matrices $\underline{C}$, we find that the set $\{\underline{D}^{(\lambda)}(B)\}$ is the full matrix algebra. The full matrix algebra being by definition irreducible, it follows that the set $\{\underline{D}^{(\lambda)}(B)\}$ constitutes an irreducible matrix representation of A_n^c.

As a last step we show which spaces carry these irreps of A_n^c. The matrix $\underline{B}$ of $B \in A_n^c$ is defined by:

$$Bu_{k\ell}^{(\lambda)} = \sum_{\mu=1}^{p} \sum_{i=1}^{f^{(\lambda)}} \sum_{j=1}^{n^{(\lambda)}} B_{ij;k\ell}^{(\mu;\lambda)} u_{ij}^{(\mu)}.$$

Inserting the expression just found for B:

$$Bu_{k\ell}^{(\lambda)} = \sum_{i=1}^{f^{(\lambda)}} D^{(\lambda)}(B)_{ik} u_{i\ell}^{(\lambda)}$$

we see that the space $W_\ell^{(\lambda)}$ spanned by $\{u_{i\ell}^{(\lambda)} \mid i = 1, \ldots, f^{(\lambda)}\}$ carries the $f^{(\lambda)}$-dimensional irreducible matrix representation $\underline{D}^{(\lambda)}(B)$ of A_n^c.

To clarify the meaning of this theorem somewhat further we write the basis elements of V_n belonging to one fixed λ in a rectangular scheme,

following ref. 15.

$$\begin{array}{ccccc} u_{11}^{(\lambda)} & u_{12}^{(\lambda)} & \cdots\cdots\cdots\cdots & & u_{1n(\lambda)}^{(\lambda)} \\ u_{21}^{(\lambda)} & \cdots\cdots\cdots & & & \vdots \\ \vdots & & & & \vdots \\ u_{f(\lambda)1}^{(\lambda)} & & \cdots\cdots\cdots\cdots\cdots\cdots & & u_{f(\lambda)n(\lambda)}^{(\lambda)} \end{array}$$

The rows in this scheme span the A_n-irreducible spaces $U_j^{(\lambda)}$ occurring in the decomposition (1) of V_n, and the columns span the A_n^c-irreducible spaces $W_i^{(\lambda)}$ occurring in the decomposition (2) of V_n.

Corollary.

The commutator algebra A_n^c of a semi-simple operator algebra A_n is semi-simple:

Proof. The theorem shows that A_n^c is decomposable; we have seen earlier that a decomposable operator algebra is semi-simple.

Corollary.

For a semi-simple operator algebra A_n the property of being a commutator algebra is symmetric, or:

$$(A_n^c)^c \equiv A_n .$$

Proof. If V_n is decomposed under A_n, it is also decomposed under A_n^c (by the foregoing theorem). Copying the proof of the foregoing theorem with A_n^c taking the place of A_n, we find that the algebra $(A_n^c)^c$ is represented by all matrices of the form: $\underline{E}_p \otimes \underline{E}_{f(\lambda)} \otimes \underline{D}^{(\lambda)}(A)$, $A \in (A_n^c)^c$. However, A_n is semi-simple, and so it contains all the matrices of this form (by Wedderburn's theorem). In other words A_n and $(A_n^c)^c$ coincide.

As we will see in sec. I.4.7, this corollary has the interesting consequence that the Casimir operators of GL(n) and U(n) are linear combinations of permutation operators.

The group theoretical approach to many body theory depends extensively on the representation theory of the Lie group GL(n) and its Lie subgroups. This family of linear Lie groups has been christened Classical Groups by Weyl [7]. Besides GL(n), the classical groups U(n) and SU(n) always play a rôle in systems of indistinguishable particles, irrespective of whether the system has geometrical symmetry or not.

In this chapter these three groups and their Lie algebras are presented, but first some of the necessary Lie group theory will be briefly outlined. For more details on this the books of Miller [13] and of Hausner and Schwartz [16] are especially recommendable.

3.1. A review of some Lie group theory

Definition.

Consider an open subset V of the r-dimensional Euclidean space F^r; the field F is either the complex field C, or the real field R. Let the zero vector $\vec{0} = (0, 0, \ldots, 0)$ belong to V. Let $\underline{A}: \vec{g} \to \underline{A}(\vec{g})$ be a mapping from V into the full matrix algebra AL(n,C), where $\underline{A}(\vec{g})$ is a matrix depending on the r parameters $\vec{g} \equiv (g_1, \ldots, g_r)$. Let V be such that the $\underline{A}(\vec{g})$ form a group; more precisely the set $G = \{\underline{A}(\vec{g}) \mid \vec{g} \in V\}$ is an r parameter linear Lie group, if it satisfies the following postulates:

G1. $\underline{A}(\vec{0}) = \underline{E}$ (the $n \times n$ unit matrix).

G2. G is a group with matrix multiplication as the group multiplication. (So, the matrices in G must be non-singular).

G3. The matrix elements $A(\vec{g})_{ij}$, $i, j = 1, \ldots, n$, are analytic functions of the r parameters $g_1, g_2, \ldots, g_r$.

G4. The r parameters are essential; this is expressed by requiring the $n^2 \times r$ Jacobi matrix $\underline{J}$ with elements:

$$J_{k\ell} \equiv \frac{\partial A_k}{\partial g_\ell}, \qquad k \equiv (i,j) = 1, \ldots, n^2; \ \ell = 1, \ldots, r$$

to have rank r. (Hence $r \leq n^2$).

Notes.

1. Postulate G4 implies immediately that the r matrices (the columns of $\underline{J}$): $\partial\underline{A}/\partial g_1$, $\partial\underline{A}/\partial g_2$, ..., $\partial\underline{A}/\partial g_r$ are linearly independent. This statement

can be turned around: if these r matrices are linearly independent, $\underline{J}$ has rank r and the r parameters are essential.

2. If the field F is the complex field, the Lie group is a complex Lie group. If F is the field R, G is a real Lie group. Notice, however, that in both cases the matrices themselves, which constitute the group, can be complex.
3. Often different mappings
$$\underline{A}: V \to G \subseteq AL(n,C)$$
of different parameter spaces V onto the linear Lie group G can be constructed. These are different parametrizations of G.
4. The definition of a linear Lie group given above is usually for so-called local linear Lie groups. For global Lie groups it is not always possible to parametrize the whole group with one open set $V \subset F^r$, often more (overlapping) subsets of F^r are required.

Examples.

1. The full rotation group SO(3), consisting of all orthogonal 3×3 matrices with determinant +1, is a Lie group. The elements of SO(3) represent rotations of R^3. Using the parametrization of Euler [17] to describe these rotations, one finds that the matrix elements contain only sines and cosines of the angles α, β and γ, and hence are analytic functions of the parameters. By differentiating an arbitrary rotation matrix $\underline{R}$ with respect to α, β and γ it is easily proved that the Euler angles are essential, for the three matrices: $\partial\underline{R}/\partial\alpha$, $\partial\underline{R}/\partial\beta$ and $\partial\underline{R}/\partial\gamma$ are linearly independent. The parameter space V is given by $0 \leq \alpha < 2\pi$, $0 \leq \beta < \pi$, and $0 \leq \gamma < 2\pi$.
2. Other examples of Lie groups are GL(n), U(n) and SU(n). These groups will be considered in more detail below.

Let us next define the Lie algebra of a Lie group. Abstract Lie algebras have already been introduced in sec. I.2.1. One may wonder why Lie algebras are considered in this work, because the irreducible representations of the classical groups just mentioned will be computed by purely "global" means; that is, we will be dealing with the groups themselves rather than with their Lie algebras. But Lie algebras, besides being useful tools for obtaining representations, usually also possess physical meaning. In particular this is true for the elements of the algebras of

Lie symmetry groups: these are constants of motion. For instance, as we will see later, the Pauli spin matrices are elements of the Lie algebra of SU(2); it is just this fact that ties electron spin to Lie group theory. So we must discuss the Lie algebra of SU(2) anyway, and it does not require much extra effort to keep the discussion fairly general. One is rewarded, not only by a much broader view of the subject, but also by a much shorter proof of the fact that GL(n) and U(n) share many of their irreps (Weyl's Unitarian Trick).

Definitions.

1. Let $N(0) \equiv (-t_0, +t_0)$ be an interval on the real line, and let $G = \{\underline{A}'(\vec{g})\}$ be a Lie group of $n \times n$ matrices. The mapping $\underline{A}$ from $N(0)$ into G:

$$\underline{A}: \quad N(0) \xrightarrow{\vec{g}} V \xrightarrow{\underline{A}'} G$$

is an analytic curve in G, provided the components of $\vec{g}(t) \in V$, $t \in N(0)$, are analytic functions of t. Clearly the matrix elements of $\underline{A}(t) \equiv \underline{A}'(\vec{g}(t))$ are then also analytic in t.

In the sequel only curves with $\underline{A}(0) = \underline{E}$, i.e. curves passing through $\underline{E}$, will be considered.

2. An analytic curve in G through $\underline{E}$ has a tangent vector $\underline{a}$ at $\underline{E}$ defined by:

$$\underline{a} = \left.\frac{d\underline{A}(t)}{dt}\right|_{t=0}$$

Notes.

1. The name tangent vector may be somewhat confusing; a tangent vector is of course still an $n \times n$ matrix, not a vector in the ordinary sense.
2. A tangent vector represents a linear operator on an n-dimensional space; this operator is called an infinitesimal operator of the Lie group G.

Theorem.

The set $\Lambda(G)$ of all tangent vectors at $\underline{E}$ of all analytic curves through $\underline{E}$ in an r parameter Lie group G is a Lie algebra of dimension r.

Proof. Let $\underline{a}$ be a tangent vector at $\underline{E}$ of the analytic curve $\underline{A}(t)$. Let $\underline{b}$ be the same for $\underline{B}(t)$. Choose two arbitrary scalars λ and μ, and take t so small that $\underline{A}(\lambda t)$ and $\underline{B}(\mu t)$ are defined. Because G is a Lie group, it follows that:

$$\underline{C}(t) \equiv \underline{A}(\lambda t)\ \underline{B}(\mu t)$$

is an analytic curve in G. The tangent vector $\underline{c}$ of $\underline{C}(t)$ belongs by definition to $\Lambda(G)$. Differentiation of the definition equation of $\underline{C}(t)$ gives:

$$\underline{c} = \lambda \underline{a} + \mu \underline{b}.$$

Hence $\Lambda(G)$ is a vector space.

Let $\underline{A}(t)$ and $\underline{B}(s)$ be analytic curves in G, then:

$$\underline{C}(s,t) \equiv \underline{B}(s)\ \underline{A}(t)\ \underline{B}(s)^{-1}$$

is an analytic curve in G, if we keep s fixed. Therefore its tangent vector:

$$\underline{c}(s) \equiv \left.\frac{d\underline{C}(s,t)}{dt}\right|_{t=0} = \underline{B}(s)\ \underline{a}\ \underline{B}(s)^{-1}$$

belongs by definition to $\Lambda(G)$. Also

$$1/s[\underline{c}(s) - \underline{c}(0)] = 1/s[\underline{B}(s)\ \underline{a}\ \underline{B}(s)^{-1} - \underline{B}(0)\ \underline{a}\ \underline{B}(0)^{-1}]$$

belongs to $\Lambda(G)$, because $\underline{B}(0) = \underline{E}$ and $\Lambda(G)$ is a vector space. Therefore, using $d\underline{B}(s)/ds = -d\underline{B}(s)^{-1}/ds$, we find that:

$$\lim_{s\to 0} 1/s[\underline{c}(s) - \underline{c}(0)] = \underline{b}\,\underline{a} - \underline{a}\,\underline{b} \equiv [\underline{b}, \underline{a}]$$

belongs also to $\Lambda(G)$. So, $\Lambda(G)$ is closed under Lie multiplication.

So, this theorem states that a unique Lie algebra is associated with every Lie group. Conversely, one can ask if every Lie algebra corresponds to a Lie group. Indeed this is true: if Λ is a matrix Lie algebra, then the multiplicative closure of the set:

$$\{e^{\underline{a}} | \underline{a} \in \Lambda\}, \qquad e^{\underline{a}} \equiv \sum_{n=0}^{\infty} \frac{1}{n!} (\underline{a})^n$$

is a Lie group with Λ as its algebra. But this correspondence is <u>not</u> unique, there can be other groups that also have Λ as their Lie algebra. All the Lie groups with the same Lie algebra are homomorphic images of another group: <u>the universal covering group $\tilde{G}$</u>, which is the "largest" group with Λ as its Lie algebra. It goes beyond the scope of this survey to say more about this, but we will give later in this chapter an example of a Lie group and its covering group.

<u>Example.</u>

Define the following analytic curves through $\underline{E}$ in SO(3):

$$\underline{R}_x(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}, \quad \underline{R}_y(t) = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}, \quad \underline{R}_z(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with corresponding tangent vectors at $\underline{E}$:

$$\underline{I}_x \equiv \left.\frac{d\underline{R}_x}{dt}\right|_{t=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \underline{I}_y \equiv \left.\frac{d\underline{R}_y}{dt}\right|_{t=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \underline{I}_z \equiv \left.\frac{d\underline{R}_z}{dt}\right|_{t=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

In sec. I.2.1 we have shown that $\{\underline{L}_x \equiv i\underline{I}_x, \underline{L}_y \equiv i\underline{I}_y, \underline{L}_z \equiv i\underline{I}_z\}$ spans a 3-dimensional Lie algebra. We now see that that algebra is (i times) the Lie algebra of SO(3).

It can easily be proved by expanding the following expressions in Taylor series:

$$e^{t\underline{I}_x}, \; e^{t\underline{I}_y}, \; e^{t\underline{I}_z}$$

and by comparing with the Taylor expansions of sin t and cos t, that these exponentials are another way of writing the matrices $\underline{R}_x(t)$, $\underline{R}_y(t)$ and $\underline{R}_z(t)$ respectively. So this example illustrates that by differentiation of the analytic curves at $\underline{E}$ we go from the Lie group to its algebra, and by exponentiation we go back from the algebra to the group.

We end this synopsis of Lie group theory with a few remarks on the representations of Lie groups and their algebras. Let the Lie group G have the representation $\underline{D}$ (see ref. 13, p. 186 for the definition). The set $\underline{D}(G)$ is also a Lie group, and has an algebra consisting of the tangent vectors:

$$\underline{a} = \left.\frac{d\underline{D}(g(t))}{dt}\right|_{t=0}, \qquad g(t) \in G.$$

It can be shown that this algebra is always a Lie algebra representation of $\Lambda(G)$, the Lie algebra of G, and hence every representation of the group yields a representation of the algebra. If the representation of the group is decomposable (irreducible), then the representation of the algebra is also decomposable (irreducible).

Conversely, let $\underline{D}$ be a representation of $\Lambda(G)$. Now, the multiplicative closure of the set

$$\{e^{\underline{a}} | \underline{a} \in \underline{D}(\Lambda(G))\}$$

may or may not be a single-valued representation of G. The representation thus obtained can be multiple-valued and is then, strictly speaking, not a representation for G. However, it will always constitute a single-valued representation for the universal covering group $\tilde{G}$ of G, and therefore one often considers first $\tilde{G}$ rather than G. If the representation $\underline{D}$ of $\Lambda(G)$ is

irreducible, then so is the corresponding representation of $\tilde{G}$. Thus, the problem of computing all irreps of G can be solved by computing all the irreps of $\Lambda(G)$, and herewith of $\tilde{G}$, provided we check every representation of $\tilde{G}$ to see that it actually determines a single-valued representation of G.

Note.
At the end of this chapter we will meet SU(2) as an example of a universal covering group (of SO(3)). There is a 2-1 homomorphism from SU(2) onto SO(3). This explains the existence of double-valued representations ("double groups") of SO(3) and its subgroups.

3 2. The general linear group GL(n) and the unitary group U(n)

In this section the general linear group GL(n) and its unitary subgroup U(n) are defined together with their Lie algebras.

Definition.
The general linear group GL(n) is the set of all non-singular linear operators on an n-dimensional vector space V_n defined over C. In the sequel this group will not be distinguished from the group consisting of all complex non-singular n × n matrices.

Note.
GL(n) is a subset of the associative algebra AL(n), defined in sec. I.2.1, which has a basis $\{\underline{E}^{ij} | i, j = 1, \ldots, n\}$ introduced in sec. I.2.2. Endowing AL(n) with the Lie product:

$$[\underline{a}, \underline{b}] \equiv \underline{a}\,\underline{b} - \underline{b}\,\underline{a}, \qquad \underline{a}, \underline{b} \in AL(n)$$

this algebra becomes also a Lie algebra. Its Lie algebra structure constants are given by:

$$[\underline{E}^{ij}, \underline{E}^{k\ell}] = \delta^{jk}\,\underline{E}^{i\ell} - \delta^{\ell i}\,\underline{E}^{kj}.$$

Theorem.

(i) GL(n) is an n^2 parameter complex Lie group.
(ii) AL(n) is the Lie algebra of GL(n).

Proof.

(i) Parametrize GL(n) in the following way:

$$\underline{A} = \underline{E} + \underline{A}' \qquad \text{with } \underline{A}' = \underline{A} - \underline{E}.$$

The set of n^2 matrix elements $\{A'_{ij}\}$ of $\underline{A}'$ is a vector in C^{n^2}, the parameter space of GL(n). Now GL(n) is a Lie group, for:

G1. Set $A'_{ij} = 0$, $i, j = 1, \ldots, n \Rightarrow \underline{A} = \underline{E}$.

G2. GL(n) is a group in the algebraic sense.

G3. The matrix elements are linear, hence analytic, functions of the parameters.

G4. The n^2 parameters are essential. (If they were not, it would imply that a relation between the matrix elements of $\underline{A}$ would exist, in contradiction with the definition of GL(n)).

(ii) Every element $\underline{a} \in AL(n)$ is a tangent vector at $\underline{E}$ of the analytic curve $e^{t\underline{a}}$, which, of course, is also an $n \times n$ matrix. Because $e^{t\underline{a}}$ has the inverse $e^{-t\underline{a}}$, it belongs to GL(n) (see ref. 13, p. 159 for a proof of $(e^{t\underline{a}})(e^{-t\underline{a}}) = \underline{E}$).

To introduce U(n) and its Lie algebra, the following two lemmas are needed.

Lemma 1.

The set Λ of all skew-Hermitean $n \times n$ matrices is a real Lie algebra of dimension n^2.

Proof. Write the complex matrix element a_{ij} of $\underline{a} \in \Lambda$ as:

$$a_{ij} = \mathrm{Re}(a_{ij}) + i\,\mathrm{Im}(a_{ij}).$$

It is then easy to show that every element $\underline{a} \in \Lambda$ can be written as follows:

$$\underline{a} = \sum_{i<j} \mathrm{Re}(a_{ij})(\underline{E}^{ij} - \underline{E}^{ji}) + i \sum_{i>j} \mathrm{Im}(a_{ij})(\underline{E}^{ij} + \underline{E}^{ji}) + i \sum_{i} \mathrm{Im}(a_{ii})\underline{E}^{ii}$$

and so the n^2 matrices:

$$\underline{a}^{ij} \equiv \tfrac{1}{2}(\underline{E}^{ij} - \underline{E}^{ji}), \qquad \text{for } i < j$$

$$\underline{a}^{ij} \equiv \tfrac{1}{2}i(\underline{E}^{ij} + \underline{E}^{ji}), \qquad \text{for } i \geq j$$

constitute an n^2-dimensional basis of Λ. The expansion coefficients of $\underline{a}$ with respect to this basis are real: Λ is a real vector space.

From $[\underline{a}, \underline{a}']^{\dagger} = -[\underline{a}, \underline{a}']$, $\underline{a}, \underline{a}' \in \Lambda$, follows then that the set Λ is closed under Lie multiplication.

Note.

One can express a skew-Hermitean matrix $\underline{a}$ also directly in terms of $\{\underline{E}^{ij}\}$. Then its components with respect to this basis obey $a^{*}_{ji} = -a_{ij}$.

Lemma 2.

Any unitary n × n matrix $\underline{U}$ can be written as:

$$\underline{U} = e^{\underline{a}}$$

with $\underline{a} \in \Lambda$.

Proof. Ref. 18, p. 55.

Note.

The matrix $\underline{a}$ is not uniquely determined by $\underline{U}$.

In the following theorem no distinction will be made between linear operators and matrices.

Theorem.

(i) The set U(n), which consists of all unitary operators on V_n, is an n^2 parameter real Lie group.

(ii) The Lie algebra $\Lambda(U(n))$ of U(n) is the algebra Λ of all skew-Hermitean n × n matrices.

Proof.

(i) According to the lemmas 1 and 2, every element $\underline{U} \in U(n)$ can be expressed in n^2 real parameters x_{ij}:

$$\underline{U} = \exp\Big[\sum_{i,j} x_{ij}\, \underline{a}^{ij}\Big], \qquad x_{ij} \in R,\ \underline{a}^{ij} \in \Lambda.$$

Now:

G1. Set $x_{ij} = 0$, $i, j = 1, \ldots, n \Rightarrow \underline{U} = \underline{E}$.

G2. U(n) is a group in the algebraic sense.

G3. Expansion of the exponential function shows that the elements of $\underline{U}$ are analytic functions of the x_{ij}.

G4. The n^2 matrices:

$$\frac{\partial \underline{U}}{\partial x_{ij}} = \underline{a}^{ij}\, \underline{U}, \qquad i, j = 1, \ldots, n$$

are linearly independent, for suppose that:

$$\sum_{i,j} C_{ij} \frac{\partial \underline{U}}{\partial x_{ij}} = \Big(\sum_{i,j} C_{ij}\, \underline{a}^{ij}\Big)\, \underline{U} = 0.$$

Multiply with $\underline{U}^{-1}$ and it follows that $C_{ij} = 0$, $i, j = 1, \ldots, n$, because the matrices $\underline{a}^{ij}$ are linearly independent. (See e.g. ref. 13, p. 158 for the differentiation of exponential matrix functions).

(ii) a. An analytic curve $\underline{U}(t)$ through $\underline{E}$ obeys:

$$\underline{U}(t)^{\dagger}\underline{U}(t) = \underline{E}$$

and so, using $\underline{U}(0) = \underline{E}$:

$$\left[\frac{d\underline{U}(t)}{dt}\right]^{\dagger}_{t=0} + \left[\frac{d\underline{U}(t)}{dt}\right]_{t=0} = \underline{0}.$$

Hence the tangent vector of $\underline{U}(t)$ at $\underline{E}$ is skew-Hermitean.

So: $\Lambda(U(n)) \subseteq \Lambda$.

b. The arbitrary matrix $\underline{a} \in \Lambda$ is a tangent vector of the curve $e^{t\underline{a}}$ in $U(n)$ and so $\Lambda \subseteq \Lambda(U(n))$.

Conclusion: $\Lambda \equiv \Lambda(U(n))$.

Notes.

1. In ref 19, chapter 2 it is shown that any $n \times n$ unitary matrix can be factorized into a product of $\frac{1}{2}n(n-1)$ complex 2×2 rotations, each containing two angles, and a diagonal matrix containing n phase factors. Of the n^2 angles thus arising n are in the closed and bounded interval $[-\pi, +\pi]$, and $n(n-1)$ are in the closed and bounded interval $[-\frac{1}{2}\pi, +\frac{1}{2}\pi]$. Hence, by the Heine-Borel theorem [20, p. 35], $U(n)$ is compact. Because integrals of analytic functions over a compact set are well-defined, many of the results known from the representation theory of finite groups also hold for compact groups. One has only to replace the summations over group elements by integrations over the parameter space of the compact group. In particular it must be noted that Maschke's theorem holds for compact groups, and so we are assured that all analytic representations of $U(n)$ are decomposable.
2. The basis $\{\underline{a}^{ij}\}$ of the real Lie algebra $\Lambda(U(n))$ is obtained by a non-singular transformation of the basis $\{\underline{E}^{ij}\}$ of the complex Lie algebra $AL(n)$, hence $\{\underline{a}^{ij}\}$ can also serve as a basis for $AL(n)$. If one extends the field R, over which $\Lambda(U(n))$ has been defined, to the complex field C, then $\Lambda(U(n))$ becomes $AL(n)$. So, $AL(n)$ is the complexification of $\Lambda(U(n))$. (We have met this case already in an example in sec. I.2.1).

3 3 Connection between the irreps of GL(n) and U(n)

The group GL(n) with its unbounded parameter space C^{n^2} is obviously non-compact, and yet many of its representations are decomposable. This is no coincidence, but follows from the fact that the complex Lie group GL(n) shares its irreps with U(n); that is: the irreps of GL(n) do not decompose upon subduction to U(n), and conversely any irrep of U(n) extends to an irrep of GL(n).

Before discussing this further we prove a theorem establishing the same fact for the corresponding Lie algebras. We need the following lemma in the proof.

Lemma.

Let $\underline{A}$ be a skew-Hermitean n × n matrix and $\underline{B}$ an arbitrary n × n matrix. If $\mathrm{Tr}(\underline{A}\,\underline{B}) = \underline{0}$ it follows that $\underline{B} = \underline{0}$.

Proof. Ref. 8, p. 165.

Theorem.

(i) An irrep of the Lie algebra AL(n) stays irreducible under restriction to $\Lambda(U(n))$.

(ii) An irreducible representation of $\Lambda(U(n))$ stays irreducible under extension to AL(n), the complexification of $\Lambda(U(n))$.

Proof.

(i) Let $\underline{D}$ be an irreducible linear matrix representation of AL(n), and assume that $\underline{D}$, possibly after a similarity transformation, reduces under the restriction to $\Lambda(U(n))$. This implies that $\underline{D}(\underline{a})$ has one or more matrix elements $D(\underline{a})_{pq}$ that are equal to zero for all $\underline{a} \in \Lambda(U(n))$. Expand $\underline{a}$ in the basis $\{\underline{E}^{ij}\}$.

$$\underline{a} = \sum_{i,j=1}^{n} a_{ij}\, \underline{E}^{ij}$$

and so:

$$D(\underline{a})_{pq} = \sum_{i,j=1}^{n} a_{ij}\, D(\underline{E}^{ij})_{pq} = 0$$

with the condition $a^{*}_{ij} = -a_{ji}$. Writing $D(\underline{E}^{ij})_{pq} = b_{ji}$, we get:

$$\sum_{i,j}^{n} a_{ij}\, b_{ji} = \mathrm{Tr}(\underline{a}\,\underline{b}) = 0$$

and from the lemma $\underline{b} = \underline{0}$. This means that the matrix elements $D(\underline{E}^{ij})_{pq}$

are zero for the whole basis of AL(n), which in turn implies that $\underline{D}$ has the same blocked form for AL(n) as for $\Lambda(U(n))$. Contradiction!

(ii) Let $\underline{D}$ be an irreducible linear matrix representation of $\Lambda(U(n))$. We have shown earlier that $\underline{a} \in \Lambda(U(n))$ can be written thus:

$$\underline{a} = \sum_{i,j} a_{ij} \underline{a}^{ij}, \qquad a_{ij} \in R$$

and so:

$$\underline{D}(\underline{a}) = \sum_{i,j} a_{ij} \underline{D}(\underline{a}^{ij}), \quad a_{ij} \in R.$$

Allowing the coefficients $\{a_{ij}\}$ to become complex, the representation $\underline{D}$ extends to a representation of AL(n). If $\underline{D}$ became reducible upon this extension, it would be so for all elements of AL(n), in particular for the elements of $\Lambda(U(n))$. Contradiction!

Returning to the Lie groups, we note that not all representations of AL(n) and $\Lambda(U(n))$ yield after exponentiation single-valued representations of their respective Lie groups. It can, however, be proved [13, p. 330] that if a representation ρ of AL(n) does yield a single-valued representation D of GL(n), then the restriction ρ' of ρ to $\Lambda(U(n))$ will give upon exponentiation a single-valued representation D' of U(n). It is further almost trivial to demonstrate that D' is then the restriction of D to U(n). Now the converse holds also: if ρ' of $\Lambda(U(n))$ yields the single-valued representation D' of U(n), then D' extends to D of GL(n); D corresponds to ρ, which in turn is the complex extension of ρ'. Since we have seen in sec. I.3.1 that all irreducible Lie group representations arise from irreducible Lie algebra representations, it follows from these considerations and the theorem above that any irrep D of GL(n) stays irreducible under subduction to U(n) and conversely that any irrep D' of U(n) extends to an irrep of GL(n). So, the irreps of the non-compact complex Lie group GL(n) have all the usual properties of compact Lie groups.

Notes.

1. This procedure of looking at a unitary and hence compact subgroup of a non-compact Lie group rather than at the Lie group itself, is called by Weyl [7] the Unitarian Trick.
2. One can consider GL(n) also as a $2n^2$ parameter real Lie group, with a $2n^2$-dimensional real Lie algebra. Proceeding this way one can obtain non-decomposable irreps of GL(n). These do not have any bearing on this work, because we will always consider GL(n) as a complex Lie group.

3 4. The special unitary group SU(n)

Lemma 1.

The set Λ of traceless skew-Hermitean matrices is an (n^2-1)-dimensional real Lie algebra.

Proof. By definition Λ is a subset of $\Lambda(U(n))$, which has the basis $\{\underline{a}^{ij}\}$ introduced in sec. I.3.2. So any $\underline{a} \in \Lambda$ can be written thus:

$$\underline{a} = \sum_{i,j} a_{ij}\, \underline{a}^{ij}, \qquad a_{ij} \in R.$$

The condition $\mathrm{Tr}(\underline{a}) = 0$ then reads $\sum_i a_{ii} = 0$. This linear relation on the components of the elements in Λ defines an (n^2-1)-dimensional linear subspace in Λ, which can easily be shown to be closed under Lie multiplication.

Lemma 2.

Any unitary matrix $\underline{U}$, with $\det(\underline{U}) = +1$, can be written as:

$$\underline{U} = e^{\underline{a}}$$

with $\underline{a} \in \Lambda$.

Proof. Write $\underline{U} = e^{\underline{a}}$ (lemma 2, sec. I.3.2) and from $\det(\underline{U}) = e^{\mathrm{Tr}(\underline{a})}$ (e.g. ref. 13, p. 156) follows $\mathrm{Tr}(\underline{a}) = 0$.

Theorem.

(i) The set: $SU(n) = \{\underline{U} \mid \underline{U} \in U(n), \det(\underline{U}) = +1\}$ is an (n^2-1) parameter real Lie group.

(ii) The Lie algebra $\Lambda(SU(n))$ of $SU(n)$ is the algebra Λ of all traceless skew-Hermitean $n \times n$ matrices.

Proof.

(i) Define the following basis for Λ in terms of the basis $\{\underline{a}^{ij}\}$ of $\Lambda(U(n))$:

$$\underline{s}^{ij} = \underline{a}^{ij} - \delta^{ij}/n \sum_{i=1}^{n} \underline{a}^{ii}, \qquad i, j = 1, \ldots, n.$$

Clearly:

$$\sum_{i=1}^{n} \underline{s}^{ii} = \underline{0}$$

and so the basis is (n^2-1)-dimensional.

Write: $\underline{U} = \exp[\sum_{i,j} x_{ij} \underline{s}^{ij}]$, $x_{ij} \in R$.

Check off the Lie group axioms G1 through G4:

G1, G2, G3 as for U(n).

G4. The n^2 matrices:

$$\frac{\partial \underline{U}}{\partial x_{ij}} = \underline{s}^{ij}\,\underline{U}, \qquad i, j = 1, \ldots, n$$

obey one linear relation: $\sum_{i=1}^{n} \underline{s}^{ii}\,\underline{U} = \underline{0}$, and so there are (n^2-1) linearly independent ones among them: SU(n) has n^2-1 essential parameters.

(ii) a. From lemma 2 it follows immediately that any curve $\underline{U}(t)$ through $\underline{E}$ in SU(n) can be written as:

$\underline{U}(t) = e^{\underline{a}(t)}$, $\quad \underline{a}(t) \in \Lambda$ and we require $\underline{a}(0) = \underline{0}$.

From this:

$$\left[\frac{d\underline{U}(t)}{dt}\right]_{t=0} = \left[\underline{U}(t)\,\frac{d\underline{a}(t)}{dt}\right]_{t=0} = \left[\frac{d\underline{a}(t)}{dt}\right]_{t=0}$$

and since:

$$\mathrm{Tr}\left[\frac{d\underline{a}(t)}{dt}\right]_{t=0} = \left[\frac{d}{dt}\,\mathrm{Tr}(\underline{a}(t))\right]_{t=0} = \frac{d}{dt}\,(0) = 0$$

it follows that $\left[\frac{d\underline{U}(t)}{dt}\right]_{t=0} \in \Lambda \Rightarrow \Lambda(SU(n)) \subset \Lambda$.

b. One proves in the same way as for U(n): $\Lambda \subset \Lambda(SU(n))$.

Example.

We will outline some properties of SU(2), a group which is of great importance for electron systems. (This will be discussed in ch. I.6). We have just met a basis for SU(n); for the case n = 2 this basis consists of the following three matrices:

$$\underline{s}^{11} = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \tfrac{1}{2}i\underline{\sigma}_z, \quad \underline{s}^{21} = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \tfrac{1}{2}i\underline{\sigma}_x, \quad \underline{s}^{12} = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \equiv \tfrac{1}{2}i\underline{\sigma}_y.$$

(The fourth matrix $\underline{s}^{22}$ is equal to $-\underline{s}^{11}$). The matrices $\underline{\sigma}_i$, i = x, y, z, are the well-known Pauli spin matrices [21, p. 545]. They obey the commutation relations:

$$[\tfrac{1}{2}i\underline{\sigma}_i, \tfrac{1}{2}i\underline{\sigma}_j] = i \sum_{k=1}^{3} \varepsilon_{ijk}\,\tfrac{1}{2}i\underline{\sigma}_k, \qquad i, j, k = x, y, z = 1, 2, 3.$$

The constants ε_{ijk} are the Levi-Civita symbols, defined in the discussion of the Lie algebra $\Lambda(SO(3))$ of SO(3) (sec. I.2.1 and sec. I.3.1). Turning

back to these discussions, one sees that $\Lambda(SU(2))$ and $\Lambda(SO(3))$ are Lie algebra isomorphic. (Note the factors $\frac{1}{2}$ in front of the Pauli matrices. As we will see shortly, they make SU(2) a "double" rotation group).

It has been pointed out in sec. I.3.1 that a Lie algebra isomorphism does not necessarily carry through for the associated groups, that is, SU(2) is not necessarily isomorphic with SO(3). To explore this point further we define a 1-1 mapping between $\Lambda(SU(2))$ and $\Lambda(SO(3))$ by writing an element of $\Lambda(SO(3))$ as $i\vec{n}\cdot\vec{\underline{L}}$, and the corresponding element of $\Lambda(SU(2))$ as $i\vec{n}\cdot\vec{\underline{\sigma}}/2$, with in both cases the same vector $\vec{n} \in R^3$. Almost exactly as for SU(2), it can be shown that every element $\underline{R}(\vec{n}) \in SO(3)$ can be written thus:

$$\underline{R}(\vec{n}) = e^{i\vec{n}\cdot\vec{\underline{L}}}, \qquad i\vec{n}\cdot\vec{\underline{L}} \in \Lambda(SO(3)).$$

Every element $\underline{U}(\vec{n}) \in SU(2)$ can be written in the form

$$\underline{U}(\vec{n}) = e^{i\vec{n}\cdot\vec{\underline{\sigma}}/2}, \qquad i\vec{n}\cdot\vec{\underline{\sigma}}/2 \in \Lambda(SU(2)).$$

Let us now see if the map: $\underline{U}(\vec{n}) \to \underline{R}(\vec{n})$ of SU(2) onto SO(3) is possibly an isomorphism. To that end we expand, and after some manipulation we get:

$$\underline{R}(\vec{n}) = \underline{E} + (i\hat{n}\cdot\vec{\underline{L}})(\sin n) + (i\hat{n}\cdot\vec{\underline{L}})^2(1 - \cos n)$$

$$\underline{U}(\vec{n}) = \underline{E}(\cos n/2) + (i\hat{n}\cdot\vec{\underline{\sigma}})(\sin n/2),$$

where $\vec{n} = n\hat{n}$ and $|\hat{n}| = 1$ (see for a derivation of these relations ref. 22, p. 12 and p. 25 resp.). It is easy to show [23, p. 8] that the matrix $\underline{R}(\vec{n})$ represents a rotation around the axis $\hat{n}$ (= unit vector along $\vec{n}$) over an angle $n = |\vec{n}|$.

The first relation shows that SO(3) is covered completely when the range of $\vec{n}$ is restricted to a sphere of radius π around the origin in R^3; each point $\vec{n}$ of the sphere represents a rotation $\underline{R}(\vec{n})$ uniquely, except when $n = \pi$, for the two points $\pi\hat{n}$ and $-\pi\hat{n}$ represent the same rotation.

The second relation shows that SU(2) is covered completely when the range of $\vec{n}$ is restricted to a sphere of radius 2π around the origin in R^3; each point $\vec{n}$ of the sphere represents a 'rotation' $\underline{U}(\vec{n})$ uniquely, except that the whole surface of the sphere ($n = 2\pi$) represents $\underline{U} = -\underline{E}$.

We note in particular that as we go over the large sphere with radius 2π, we cover SU(2) once and SO(3) twice in such a way that the two points $n\hat{n}$ and $-(2\pi - n)\hat{n}$ correspond to the same rotation of SO(3) and to elements of SU(2) of opposite sign.

Summarizing, we have found that two different elements of SU(2) map onto the same element of SO(3):

$$\left.\begin{matrix}\underline{U}(n\hat{n}) \\ \underline{U}(-(2\pi - n)\hat{n})\end{matrix}\right\} \rightarrow \underline{R}(n\hat{n}), \qquad 0 \leq n < 2\pi.$$

From the Campbell-Baker-Hausdorff formula [13, p. 161] and the isomorphism of $\Lambda(SO(3))$ and $\Lambda(SU(2))$ follows that this map is a group homomorphism; SU(2) is a double-valued representation of SO(3).

Now, it can be proved by topological arguments that SU(2) is the "largest" group with a Lie algebra which is isomorphic to $\Lambda(SO(3))$, and so SU(2) is the universal covering group of SO(3).

The space $V_n \otimes^N$, which is the N-fold tensorial product of V_n with itself, carries representations of the permutation group S_N on the one hand and of the general linear group GL(n) on the other. In this chapter it will be shown that the representations of these two groups are closely related: decomposition of $V_n \otimes^N$ under the one group induces simultaneous decomposition under the other.

This remarkable fact was discovered by Schur in 1901 and later explored in depth by Weyl [6, 7], who also recognized its importance for quantum mechanics. Because Weyl's expositions make extremely difficult reading, and because Boerner's exposition of the same subject [8] is not much more transparent, some of the more important proofs will be presented in the hope that I have succeeded in clarifying the derivations for non-mathematical readers. At the same time the theory is extended to include also the Lie algebra of GL(n), thus leading up to a rigorous proof of the fact that the Casimir operators of GL(n) are represented by linear combinations of permutation operators. This relation between Casimir operators and permutation operators has been noted before [24, 25], but it seems that no general proof has been given earlier.

4.1. Tensor products of vector spaces

Definition.

Let U be an n-dimensional vector space with basis $\{u_i\}$ and V an m-dimensional vector space with basis $\{v_j\}$. Associate with each element (u_i, v_j) in the Cartesian product $U \times V$ an element $u_i \otimes v_j$, and let W be the nm-dimensional linear closure consisting of all formal sums:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\, u_i \otimes v_j, \qquad c_{ij} \in C.$$

If we further postulate:

T1. $(cu_i) \otimes v_j = u_i \otimes (cv_j) = c(u_i \otimes v_j), \qquad c \in C$

T2 a. $u_k \otimes (v_i + v_j) = u_k \otimes v_i + u_k \otimes v_j$

T2 b. $(u_i + u_j) \otimes v_k = u_i \otimes v_k + u_j \otimes v_k$

for all basis vectors u_i of U and v_j of V, then W is a tensor product space, commonly denoted by $U \otimes V$.

Notes.

1. From the postulates follows immediately:

$$0 \otimes v = u \otimes 0 = 0 \otimes 0 \in U \otimes V, \qquad \forall u \in U, \forall v \in V$$

and $u \otimes v = 0$ implies either: $u = 0$ or $v = 0$ or $u = v = 0$. It then follows that $U \otimes V$ is nm-dimensional.

2. A tensor product $u \otimes v \in U \otimes V$ is not commutative.

3. Nothing has been required of the vector spaces U and V except finite dimensionality, and so these spaces can be tensor product spaces themselves. Postulating:

   T3. $(U \otimes V) \otimes W = U \otimes (V \otimes W)$

   the tensor product of vector spaces can be extended to an arbitrary number of factors.

4. The N-fold tensor product $V \otimes V \otimes \ldots \otimes V$ of the vector space V with itself will in the sequel be denoted by $V \otimes^N$.

5. Elements of $V \otimes^N$ are called tensors.

6. The definition of a tensor product space is formulated in terms of bases of the vector spaces constituting the product. (See e.g. ref. 12, p. 85 for a basis free definition, which has the advantage of holding for infinite dimensional spaces as well. This latter definition will tacitly be assumed in ch. I.6, where tensor products of infinite dimensional spaces are considered).

Strictly speaking, it must now be proved that a tensor product space is independent of the choice of bases in the different vector spaces. Since this proof is a trivial generalization of the following theorem it will be forgone. Summation convention will be used in the rest of this section.

Theorem.

Let $\{v_i\}$ and $\{v'_j\}$ be bases of V_n connected via:

$$v'_j = a^i_j \, v_i .$$

Then the corresponding bases of $V_n \otimes^N$ satisfy the following transformation equation:

$$v'_{j_1} \otimes v'_{j_2} \otimes v'_{j_3} \otimes \ldots \otimes v'_{j_N} = a^{i_1}_{j_1} a^{i_2}_{j_2} a^{i_3}_{j_3} \ldots a^{i_N}_{j_N} \, v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_N} .$$

Proof. Direct consequence of the postulates.

Note.

If a^{i}_{j} is the (i,j)-element of $\underline{A}$, then $a^{i_1}_{j_1} a^{i_2}_{j_2} \ldots a^{i_N}_{j_N}$ is the $(i_1,i_2,\ldots,i_N; j_1,j_2,\ldots,j_N)$-element of the Kronecker product matrix $[\underline{A}]^N \equiv \underline{A} \otimes \underline{A} \otimes \underline{A} \otimes \ldots \otimes \underline{A}$, (see sec. I.2.3 for the definition).

Notation.

1. In the sequel we will often write E_I instead of $v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_N}$ for a basis vector of $V_n{}^{\otimes N}$; I then stands for the index set $\{i_1, i_2, \ldots, i_N\}$. The theorem just stated reads in this notation:

$$v'_j = a^i_j v_i \Rightarrow E'_J = A^I_J E_I,$$

where A^I_J is the $(I;J)$-element of $[\underline{A}]^N$.

2. A tensor $T \in V_n{}^{\otimes N}$ can be expanded with respect to the basis $\{E_I\}$:

$$T = E_I t^I.$$

Here $t^I \equiv t^{i_1 i_2 i_3 \ldots i_N}$ is the component of T along E_I.

Theorem.

Let $\{E'_J\}$ and $\{E_I\}$ be two different bases of $V_n{}^{\otimes N}$ connected via:

$$E'_J = A^I_J E_I$$

and let T have the components t'^J and t^I with respect to the basis $\{E'_J\}$ and $\{E_I\}$ respectively. Then

$$t^I = A^I_J t'^J.$$

Proof. $T = E_I t^I = E'_J t'^J = E_I A^I_J t'^J \Rightarrow t^I = A^I_J t'^J$.

Notes.

1. An array consisting of n^N scalars t^I can be taken to represent a tensor in $V_n{}^{\otimes N}$ if it behaves under a change of basis in V_n as in the theorem above. This leads to the definition of a tensor most commonly found in books on vector and tensor analysis.
2. Tensors are important for this work, because orbital products are but an example of tensors. It is common in quantum chemistry to denote tensors of this kind by particle labels rather than by tensor product symbols. Keeping track of particle labels one can commute factors in a

tensor product. For instance $\phi(1)\psi(2)$ and $\psi(2)\phi(1)$ stand both for the same tensor $\phi \otimes \psi$.

4.2. The space $V_n{}^{\otimes N}$ as a carrier space for representations of GL(n)

Definitions.

1. Let α and β be linear operators on V_n. The tensor product $\alpha \otimes \beta$ is the linear operator on $V_n{}^{\otimes^2}$ defined by:

$$\alpha \otimes \beta \ (u \otimes v) = \alpha(u) \otimes \beta(v), \qquad u, v \in V_n.$$

If α has the matrix $\underline{A}$ and β has the matrix $\underline{B}$, then $\alpha \otimes \beta$ has the matrix $\underline{A} \otimes \underline{B}$.

2. Tensor products of operators multiply thus:

$$(\alpha \otimes \beta)(\alpha' \otimes \beta') = (\alpha\alpha') \otimes (\beta\beta')$$

(c.f. the multiplication of Kronecker product matrices).

3. Tensor products of operators can easily be defined for more than two factors, postulating associativity.

Theorem.

The set of linear operators on $V_n{}^{\otimes N}$:

$$GL(n)^{\otimes N} = \{\alpha' \otimes \alpha'' \otimes \ldots \otimes \alpha^{(N)} \mid \alpha', \alpha'', \ldots, \alpha^{(N)} \in GL(n)\}$$

is a group: the outer direct product group of GL(n).

Proof. Multiplication of linear operators is associative. The identity is $e \otimes e \otimes \ldots \otimes e$, $e \in GL(n)$. The inverse of $\alpha' \otimes \alpha'' \otimes \ldots \otimes \alpha^{(N)}$ is $(\alpha')^{-1} \otimes (\alpha'')^{-1} \otimes \ldots \otimes (\alpha^{(N)})^{-1}$.

Theorem.

(i) The set of linear operators on $V_n{}^{\otimes N}$:

$$[GL(n)]^N = \{\alpha \otimes \alpha \otimes \ldots \otimes \alpha \mid \alpha \in GL(n)\}$$

is a group: the inner direct product group of GL(n).

(ii) Let C_N be the cyclic group of order N, then:

$$GL(n)/C_N \cong [GL(n)]^N.$$

Proof.

(i) In the same way as for the outer direct product of GL(n).

(ii) Define the mapping τ: $GL(n) \to [GL(n)]^N$ by

$$\tau(\alpha) = \alpha \otimes \alpha \otimes \ldots \otimes \alpha, \qquad \alpha \in GL(n),$$

which is readily seen to be a group homomorphism and an onto mapping. The kernel of τ consists of the elements α with $\tau(\alpha) \equiv \alpha \otimes \alpha \otimes \ldots \otimes \alpha = e \otimes e \otimes \ldots \otimes e$. From this follows: $\alpha = xe$ with $x^N = 1$. So the kernel of τ is the set:

$$C_N = \{x_k e \,|\, x_k = e^{(2\pi i/N)k},\ e \in GL(n)\},$$

which is a cyclic subgroup of GL(n). By theorem 15 of ref. 4, it follows that $GL(n)/C_N \cong [GL(n)]^N$.

We have now found that the inner direct product group $[GL(n)]^N$ represents GL(n) on $V_n{}^{\otimes^N}$. This representation is non-faithful. In the following example we will furthermore see that it is a reducible one.

Example.

Two well-known processes often applied to second rank tensors are symmetrization and contraction. Both procedures have a profound group theoretical meaning: the decomposition of a second rank tensor into a symmetric and an antisymmetric component adapts the tensor to GL(n), i.e. decomposes it into components belonging to irreducible carrier spaces of GL(n). Contraction of indices gives a further adaptation of the tensor to the orthogonal group $O(n) \subset GL(n)$. For more about the latter point the reader is referred to ref. 26, sec. 10-5 to 10-7.

We will presently show that the representation of GL(n) on $V_n{}^{\otimes^2}$ can be reduced by decomposing this tensor space into a direct sum of a symmetric and an antisymmetric subspace. Define a new basis of $V_n{}^{\otimes^2}$ by:

$$w^+_{ij} = v_i \otimes v_j + v_j \otimes v_i, \qquad i \leq j = 1, \ldots, n$$

$$w^-_{ij} = v_i \otimes v_j - v_j \otimes v_i, \qquad i < j = 1, \ldots, n.$$

Now:

$$\begin{aligned}
\alpha \otimes \alpha(w^+_{ij}) &= a^k_i a^\ell_j\, v_k \otimes v_\ell + a^k_j a^\ell_i\, v_k \otimes v_\ell \\
&= \tfrac{1}{2}[a^k_i a^\ell_j\, v_k \otimes v_\ell + a^\ell_i a^k_j\, v_\ell \otimes v_k + a^k_j a^\ell_i\, v_k \otimes v_\ell + a^\ell_j a^k_i\, v_\ell \otimes v_k] \\
&= \tfrac{1}{2}[a^k_i a^\ell_j + a^k_j a^\ell_i]\, v_k \otimes v_\ell + \tfrac{1}{2}[a^k_j a^\ell_i + a^k_i a^\ell_j]\, v_\ell \otimes v_k \\
&= \tfrac{1}{2}[a^k_i a^\ell_j + a^\ell_i a^k_j]\,(v_k \otimes v_\ell + v_\ell \otimes v_k) \\
&= T^+(\alpha)^{k\ell}_{ij}\, w^+_{k\ell}.
\end{aligned}$$

So, the $\frac{1}{2}n(n+1)$-dimensional subspace of $V_n^{\otimes^2}$ spanned by the symmetric tensors w^+_{ij} is invariant under $\alpha \otimes \alpha$. The matrix $T^+(\alpha)^{k\ell}_{ij} = \frac{1}{2}[a^k_i a^\ell_j + a^\ell_i a^k_j]$ is an example of an irreducible tensor representation of GL(n). In chapter I.5 general expressions for tensor representations will be derived.

For the antisymmetric tensors one gets similarly:

$$\alpha \otimes \alpha(w^-_{ij}) = \tfrac{1}{2}[a^k_i a^\ell_j - a^\ell_i a^k_j]\ (v_k \otimes v_\ell - v_\ell \otimes v_k)$$

$$= T^-(\alpha)^{k\ell}_{ij}\ w^-_{k\ell},$$

and so the $\frac{1}{2}n(n-1)$-dimensional subspace of $V_n^{\otimes^2}$ spanned by the antisymmetric tensors w^-_{ij} is also invariant under $\alpha \otimes \alpha$.

The main purpose of this chapter is twofold: the processes of symmetrization and antisymmetrization will be generalized to tensors of arbitrary rank N, and the representations of GL(n) herewith obtained will be shown to be irreducible.

4.3. The space $V_n^{\otimes^N}$ as a carrier space for representations of the Lie algebras of the classical groups

Let G be a classical group. By definition G consists of linear operators on V_n. Applying the argument of the foregoing section, one sees that the inner direct product group $[G]^N$ forms a (possibly non-faithful) representation of G, carried by $V_n^{\otimes^N}$. The Lie algebra $\Lambda([G]^N)$ of $[G]^N$ consists also of linear operators on $V_n^{\otimes^N}$. We will show that this Lie algebra is a faithful representation of $\Lambda(G)$, the Lie algebra of G.

Definition.

Let $\underline{a}$ be an arbitrary n × n matrix and $\underline{E}$ the n × n unit matrix, then we define:

$$\underline{a}(k) \equiv \underline{E} \otimes \underline{E} \otimes \ldots \underline{E} \otimes \underline{a} \otimes \underline{E} \ldots \otimes \underline{E}$$

where the matrix on the right hand side is an N-fold Kronecker product matrix having the matrix $\underline{a}$ as the k-th factor and unit matrices as the other factors.

Theorem.

(1) The Lie algebra $\Lambda([G]^N)$ consists of all possible elements of the form

$$\sum_{k=1}^{N} \underline{a}(k), \qquad \underline{a} \in \Lambda(G).$$

(ii) The mapping $\underline{a} \rightarrow \sum_{k=1}^{N} \underline{a}(k)$ of $\Lambda(G)$ onto $\Lambda([G]^N)$ is a Lie algebra isomorphism.

Proof.

(i) Let $\underline{a}$ be the tangent vector at $\underline{E}$ of the analytic curve $\underline{A}(t)$ in G, so $\underline{a}$ belongs to $\Lambda(G)$. Now $\underline{A}(t) \otimes \underline{A}(t) \otimes \ldots \otimes \underline{A}(t)$ (N factors) is obviously an analytic curve in $[G]^N$ with tangent vector:

$$\frac{d}{dt}[\underline{A}(t) \otimes \underline{A}(t) \ldots \otimes \underline{A}(t)]_{t=0} = \sum_{k=1}^{N} \underline{E} \otimes \underline{E} \ldots \otimes \underline{a} \otimes \ldots \otimes \underline{E} = \sum_{k=1}^{n} \underline{a}(k).$$

This follows immediately by application of the chain rule, which can easily be asserted for Kronecker product matrices.

(ii) The mapping of $\Lambda(G)$ onto $\Lambda([G]^N)$ is clearly one-to-one; it is also easy to show that it is linear and preserves commutator brackets (Lie products).

Subsequently, an example of the Lie algebra of an inner direct product group will be presented. In this example we will need the following lemma.

Lemma.

Let $\underline{a}$, $\underline{b}$, ..., $\underline{z}$ be m arbitrary n × n matrices. Then:

$$e^{\underline{a}} \otimes e^{\underline{b}} \otimes \ldots \otimes e^{\underline{z}} = \exp[\underline{a}(1) + \underline{b}(2) + \ldots + \underline{z}(m)].$$

Proof.

$$e^{\underline{a}} \otimes e^{\underline{b}} = \sum_{i,j} \frac{1}{i!j!} (\underline{a})^i \otimes (\underline{b})^j = \sum_{i,j} \frac{1}{i!j!} (\underline{a} \otimes \underline{E})^i (\underline{E} \otimes \underline{b})^j = e^{\underline{a} \otimes \underline{E}}\, e^{\underline{E} \otimes \underline{b}} =$$

$$= \exp[\underline{a} \otimes \underline{E} + \underline{E} \otimes \underline{b}] \equiv \exp[\underline{a}(1) + \underline{b}(2)].$$

By induction one extends this result to m factors.

Example.

An element of the classical group SU(2) can be written as:

$$U(\vec{n}) = e^{i\vec{n} \cdot \vec{s}}$$

(sec. I.3.4; here we employ the one-electron spin operator $\vec{s}$ rather than its matrix $\vec{\sigma}/2$).

An element of $[SU(2)]^N$ has the form:

$$[U(\vec{n})]^N = U(\vec{n}) \otimes U(\vec{n}) \otimes \ldots \otimes U(\vec{n})$$

and from the lemma follows:

$$[U(\vec{n})]^N = \exp[i\vec{n} \cdot \vec{S}]$$

with:

$$\vec{S} = \sum_{k=1}^{N} \vec{s}(k).$$

The components iS_x, iS_y, iS_z of $i\vec{S}$ span clearly the Lie algebra $\Lambda([SU(2)]^N)$, which is isomorphic to $\Lambda(SU(2))$. Quantum mechanically $\vec{S}\cdot\vec{S}$ is the observable representing the spin of an N-electron system.

4.4. The space $V_n{}^{\otimes N}$ as a carrier space for representations of CS_N

The tensor space $V_n{}^{\otimes N}$ carries in a very natural way representations of the permutation group S_N and its group algebra CS_N. However, different conventions may be chosen to define the action of permutation operators on tensors and since erroneous results are obtained when not strictly adhering to a convention once chosen, we must define unambiguously how a permutation acts on a tensor.

Definition.

A permutation $\pi \in S_N$ acts on a set A of N objects, which are not necessarily different. The positions (not the objects!) in A are numbered from 1 to N. The action of

$$\pi = \begin{pmatrix} 1 & 2 & \dots & k & \dots & N \\ j_1 & j_2 & \cdots & j_k & \cdots & j_N \end{pmatrix}$$

on A is described by the rule Put object standing in position k into position j_k, for k = 1, 2, ..., N successively.

Example.

From here on the well-known cycle notation will be used to denote permutations (see ref. 26, p. 13 for a definition of this notation).

Let A be the set {a b c d b} and number its positions, counting from 1 to 5 going from left to right. Now:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 5 & 1 \end{pmatrix} A \equiv (1345)A = \{b\ b\ a\ c\ d\} = A'$$

$$(123)(1345)A = (123)A' = \{a\ b\ b\ c\ d\} = (2345)A.$$

So we have the multiplication rule: (123)(1345) = (2345).

Definition.

The action of an arbitrary permutation $\pi \in S_N$ on an index set $I = \{i_1, i_2, \ldots, i_N\}$ is defined as above, after numbering the positions in I from left to right. Writing the basis of $V_n \otimes^N$ as $\{E_I \mid \{I = i_1, \ldots, i_N\}\}$ and an arbitrary tensor T with respect to this basis as $t^I E_I$, we can give two alternative definitions of a linear operator $\hat{\pi}$: $V_n \otimes^N \to V_n \otimes^N$ associated with $\pi \in S_N$:

(i) $$\hat{\pi} E_I = E_{\pi(I)},$$

(ii) $$\hat{\pi} t^I = t^{\pi^{-1}(I)}.$$

Notes.

1. It is easy to show that the two definitions of $\hat{\pi}$ give the same linear operator on $V_n \otimes^N$. For map an arbitrary tensor $T \in V_n \otimes^N$:

$$\hat{\pi} T = \hat{\pi}(t^I E_I) = t^I E_{\pi(I)} = t^{\pi^{-1}(I)} E_I = (\hat{\pi} t^I) E_I,$$

where we used the fact that I runs over all possible index sets. So we can define $\hat{\pi}$ by its action on the covariant (lower) indices or by its action on the contravariant (upper) indices of a tensor.

2. The mapping $\pi \to \hat{\pi}$ is a group homomorphism, that is to say:

(i) $\hat{\tau}\hat{\pi} E_I = \hat{\tau} E_{\pi(I)} = E_{\tau\pi(I)}$

(ii) $\hat{\tau}\hat{\pi} t^I = \hat{\pi} t^{\tau^{-1}(I)} = t^{\pi^{-1}\tau^{-1}(I)} = t^{(\tau\pi)^{-1} I}$.

Example.

In the following chapter Young operators of the NP-type will often be used. An example of such an operator is:

$$\hat{Y} = [(1) - (13)][(1) + (12)] = (1) + (12) - (13) - (123).$$

Now

$$\hat{Y} E_{121} = E_{121} + E_{211} - E_{121} - E_{112} = E_{211} - E_{112}$$

$$\hat{Y} t^{121} = t^{121} + t^{211} - t^{121} - t^{211} = 0.$$

Note that $\hat{Y}$ acting on contravariant indices gives zero, whereas acting on covariant indices it gives a result not equal to zero. Had we used a Young operator of the PN-type, we would have found the converse of this result.

Having defined representations for S_N, we can now turn to CS_N.

Definition.

The enveloping algebra of the set of permutation operators on $V_n{}^{\otimes N}$ will be denoted by AS_N. This algebra represents the group algebra CS_N on $V_n{}^{\otimes N}$

In the sequel we drop the caret over the permutation operators, not distinguishing in the notation any further between permutations (acting on arbitrary sets) and permutation operators (acting on $V_n{}^{\otimes N}$).

4.5. Bisymmetric operators on $V_n{}^{\otimes N}$

The linear operators on $V_n{}^{\otimes N}$ that belong to the commutator algebra AS_N^C of AS_N are often called bisymmetric operators. In this section it will be demonstrated that the algebra of bisymmetric operators is generated by $[N(\underline{E})]^N$, the inner direct product of the local Lie subgroup $N(\underline{E})$ of $GL(n)$. This has two consequences, viz.:

- AS_N^C is the enveloping algebra of $[GL(n)]^N$.
- AS_N^C is the enveloping algebra of $\Lambda([GL(n)]^N)$; in other words: every bisymmetric operator can be written as a polynomial in the infinitesimal generators of $[GL(n)]^N$.

Definition.

$N(\underline{E})$ is the set of all complex $n \times n$ matrices with matrix elements satisfying:

$$|a_{ij} - \delta_{ij}| < 1/n, \qquad (a_{ij}) \in N(\underline{E}).$$

Lemma 1.

The matrices in $N(\underline{E})$ are non-singular, so. $N(\underline{E}) \subset GL(n)$.

Proof. Define a matrix norm via:

$$\|\underline{A}\| = [\mathrm{Tr}(\underline{A}^{\dagger}\underline{A})]^{\frac{1}{2}}$$

(ref. 18, p. 53). Here $\underline{A}^{\dagger}$ is the Hermitean adjoint of $\underline{A}$ and $\mathrm{Tr}\underline{A}$ stands for the trace of $\underline{A}$.

We show that $\|\underline{A} - \underline{E}\|^2 < 1$ for all $\underline{A} \in N(\underline{E})$:

$$\|\underline{A} - \underline{E}\|^2 = \sum_{i,j} (a_{ji} - \delta_{ji})^{*}(a_{ij} - \delta_{ij}) = \sum_{i,j} |a_{ij} - \delta_{ij}|^2 < \sum_{i,j} 1/n^2 = 1.$$

A matrix $\underline{A}$ with $\|\underline{A} - \underline{E}\|^2 < 1$ is non-singular, for:

$\underline{A} = \underline{E} + (\underline{A} - \underline{E}) \equiv \underline{E} + \underline{A}' = e^{\ell n(\underline{E}+\underline{A}')}$ and $\underline{A}^{-1} = e^{-\ell n(\underline{E}+\underline{A}')}$.
This way of writing $\underline{A}$ and $\underline{A}^{-1}$ is permitted since $\ell n(\underline{E} + \underline{A}')$ is well-defined if $\|A'\| < 1$ (ref. 18, p. 55).

Notes.

1. $N(\underline{E})$, being a subset of GL(n) and a neighbourhood of $\underline{E}$, is a local Lie group (ref. 13, p. 162).
2. In the same way as for GL(n), one defines the inner direct product $[N(\underline{E})]^N$ of $N(\underline{E})$.

We now turn to bisymmetric operators.

Definition.
The elements of AS_N^C, the commutator algebra (sec. I.2.1) of AS_N, are called bisymmetric operators.

The name bisymmetric is due to Weyl [7] and was inspired by the following lemma.

Lemma 2.
Let the linear operator β on $V_n{}^{\otimes N}$ have the matrix (B_J^I), then:

$$B_J^I = B_{\pi(J)}^{\pi(I)}, \qquad \forall \pi \in S_N$$

if and only if $\beta \in AS_N^C$.

Proof.

(i) Assume $\beta \in AS_N^C$:

$$B_I^J E_J \equiv \beta E_I = \beta\pi^{-1}\pi E_I = \pi^{-1}\beta E_{\pi(I)} = \pi^{-1} B_{\pi(I)}^{\pi(J)} E_{\pi(J)} = B_{\pi(I)}^{\pi(J)} E_J \Rightarrow$$

$$\Rightarrow B_I^J = B_{\pi(I)}^{\pi(J)}.$$

(ii) Assume $B_J^I = B_{\pi(J)}^{\pi(I)}$:

$$\beta\pi E_I = \beta E_{\pi(I)} = B_{\pi(I)}^{\pi(J)} E_{\pi(J)} = B_I^J E_{\pi(J)} = \pi B_I^J E_J = \pi\beta E_I.$$

This lemma enables us to recognize bisymmetric operators, as in the following examples.

Examples.

1. The elements of $[GL(n)]^N$ are bisymmetric, for the (I,J)-element of $[\underline{A}]^N \in [GL(n)]^N$ is:

$$A^I_J = a^{i_1}_{j_1} a^{i_2}_{j_2} \cdots a^{i_N}_{j_N}.$$

The multiplication of the factors on the right hand side is commutative, and so $A^{\pi(I)}_{\pi(J)} = A^I_J$ for all $\pi \in S_N$.

Note that the elements of the enveloping algebra (sec. I.2.1) of $[GL(n)]^N$ are also bisymmetric; in fact we will show later in this section that this enveloping algebra is an improper subalgebra of AS^C_N, i.e. $\mathcal{E}([GL(n)]^N) = AS^C_N$.

2. The elements of the Lie algebra $\mathcal{A}([GL(n)]^N)$ are bisymmetric. The enveloping algebra of this Lie algebra is also contained in AS^C_N.
3. A Hamiltonian describing a system of N identical particles is bisymmetric. The physical consequences of this observation will be worked out in ch. I.6.

By definition the matrix of a linear operator on $V_n{}^{\otimes N}$ belongs to $AL(n^N,C)$, which is an algebra of dimension $(n^N)^2$. From the equality of dimensions follows that this full matrix algebra is equal to $AL(n,C)^{\otimes N}$, for this tensor product has dimension $(n^2)^N$.

The algebra of bisymmetric mappings is a proper subspace of $AL(n,C)^{\otimes N}$. This is the content of the following lemma.

Lemma 3.

AS^C_N is a proper subspace of $AL(n,C)^{\otimes N}$ of dimension $\binom{n^2 + N - 1}{N}$.

Proof. If one considers $\beta \in AS^C_N$ as a vector, then its matrix elements B^I_J are its components with respect to the basis of $AL(n^N,C)$.
Note that (B^I_J) is not necessarily a Kronecker product matrix, or in other words $B^I_J \equiv B^{i_1 \cdots i_N}_{j_1 \cdots j_N}$ is not necessarily a product of the form

$$b^{i_1}_{j_1} b^{i_2}_{j_2} \cdots b^{i_N}_{j_N}.$$

Recalling from linear algebra that the number of linearly independent components of an arbitrary vector is equal to the dimension of the subspace to which the vector belongs, we see that the bisymmetric operators certainly belong to a proper subspace of $AL(n^N,C)$; for the relation

$B^{\pi(I)}_{\pi(J)} = B^{I}_{J}$ shows that many of the components of β are equal and hence linearly dependent.

Restrict now the index sets labelling the components $B^{i_1 \cdots i_N}_{j_1 \cdots j_N}$ of β in the following manner:

(i) Replace the double index (i_p, j_p) by a single index $k_p = (n-1)i_p + j_p$ for all $p = 1, 2, \ldots, N$.

(ii) Consider only ordered index sets K, i.e. impose on the elements of K the condition:

$$k_1 \leq k_2 \leq \cdots \leq k_N.$$

The bisymmetry of β yields in this notation:

$$B^{\pi(K)} = B^{K}, \qquad \forall \pi \in S_N,$$

and so the set $\{B^K\}$ contains only the components of β which do not follow from each other by permutation. Any component of β not contained in the set $\{B^K\}$ is equal to one of the B^K in the set.

Since no condition other than commutation with all permutation operators has been imposed on bisymmetric operators, we have exhausted all the relations between the components of β: the elements of $\{B^K\}$ are linearly independent and so the dimension of AS^C_N is equal to the order of $\{B^K\}$.

Finally, it is known from combinatorics (see e.g. ref. 27, p. 488) that the number of ordered index sets $K = \{k_1 \leq k_2 \leq \cdots \leq k_N\}$ with $k_p = 1, 2, \ldots, n^2$ is equal to $\binom{n^2 + N - 1}{N}$.

We now come to the crucial theorem of this section. The following lemma is needed in the proof.

Lemma 4.

Let T_i, $i = 1, \ldots, n^2$, be n^2 infinite subsets of C. Let $P(x^1, x^2, \ldots, x^{n^2})$ be a polynomial in n^2 variables in C. If for each i ($i = 1, \ldots, n^2$) $P(x^1, x^2, \ldots, x^{n^2}) = 0$ for all points $x^i \in T_i$ independently of the remaining variables, then $P \equiv 0$. That is, the coefficients of P are all zero.

Proof. Ref. 28, p. 121, corollary 2.

Theorem.

The algebra AS^C_N is the enveloping algebra of $[N(\underline{E})]^N$.

Proof. Note first that the enveloping algebra $\mathcal{A}([N(\underline{E})]^N)$ of $[N(\underline{E})]^N$ is contained in AS_N^C, for we have: $[N(\underline{E})]^N \subset [GL(n)]^N \subset AS_N^C$, and products and linear combinations of bisymmetric operators are bisymmetric, so closing $[N(\underline{E})]^N$ multiplicatively and linearly does not bring us outside AS_N^C.

Note secondly that both AS_N^C and $\mathcal{A}([N(\underline{E})]^N)$ are subspaces of $AL(n,C)^{\otimes N}$. As we will show that these spaces have the same dimension, they must coincide.

To compute the dimension of $\mathcal{A}([N(\underline{E})]^N)$ we consider how many of the matrix elements of an arbitrary matrix $[\underline{X}]^N \in [N(\underline{E})]^N$ are linearly independent. Write to that end firstly the elements of $[X]^N$ thus:

$$x^{i_1}_{j_1} x^{i_2}_{j_2} \ldots x^{i_N}_{j_N} = x^{k_1} x^{k_2} \ldots x^{k_N} \equiv x^K$$

with $k_p = (n-1)i_p + j_p$, $p = 1, \ldots, N$, $k_p = 1, \ldots, n^2$. In the manner of the proof of lemma 3 it now follows immediately that only the $\binom{n^2 + N - 1}{N}$ elements x^K with ordered index set K are essentially different.

The matrix elements $x^{i_p}_{j_p}$ satisfy the condition:

$$\left| x^{i_p}_{j_p} - \delta^{i_p}_{j_p} \right| < 1/n, \qquad p = 1, \ldots, N,$$

demonstrating that the diagonal elements belong to a neighbourhood $T_1 \subset C$ of 1 and that the non-diagonal elements belong to a neighbourhood $T_0 \subset C$ of 0. Each element $x^{i_p}_{j_p} = x^{k_p}$ may be chosen independently of the remaining elements from its neighbourhood; any choice of n^2 elements yields a non-singular (by lemma 1) matrix belonging to $N(\underline{E})$.

The monomials x^K are linearly independent. To prove this we note that a linear relation among them is necessarily trivial (i.e. has all coefficients equal to zero), because the following linear relation between the essentially different matrix elements x^K of $[\underline{X}]^N$:

$$P(x^1, \ldots, x^{n^2}) \equiv \sum_{k_1 \leq k_2 \cdots \leq k_N} C_{k_1 k_2 \ldots k_N} x^{k_1} x^{k_2} \ldots x^{k_N} = 0$$

implies by lemma 4 that all $C_{k_1 \ldots k_N}$ are zero.

Summarizing: an arbitrary matrix $[\underline{X}]^N \in [N(\underline{E})]^N$ has $\binom{n^2 + N - 1}{N}$ linearly independent matrix elements (the components of $[\underline{X}]^N$ with respect to the basis of $AL(n,C)^{\otimes N}$), hence $[N(\underline{E})]^N$ generates an $\binom{n^2 + N - 1}{N}$-dimensional subspace (the algebra of bisymmetric operators) of $AL(n,C)^{\otimes N}$.

Note.

In the proof just given it is essential that each of the elements x^i $(i = 1, \ldots, n^2)$ can be chosen independently of the remaining elements. If this is not the case, lemma 4 is not applicable. Therefore, the same proof does not hold for Lie groups, consisting of n × n matrices, that have less than n^2 parameters.

Corollary.

The algebra AS_N^C is the enveloping algebra of $[GL(n)]^N$.

Proof. The enveloping algebra of $[GL(n)]^N$ is obviously contained in AS_N^C. The subset $[N(\underline{E})]^N$ of $[GL(n)]^N$ already generates all of AS_N^C and hence $[GL(n)]^N$ does so a fortiori.

Theorem.

The algebra AS_N^C is the enveloping algebra of $\Lambda([GL(n)]^N)$.

Proof. The enveloping algebra of $\Lambda([GL(n)]^N)$ is obviously contained in AS_N^C. We will show that $[N(\underline{E})]^N$ is a subset of this enveloping algebra, from which the theorem follows immediately.
The matrix $\ell n([\underline{X}]^N)$, $[\underline{X}]^N \in [N(\underline{E})]^N$, is well-defined and belongs to $\Lambda([GL(n)]^N)$:

$$\begin{aligned}
\ell n([\underline{X}]^N) &= \ell n(\underline{X} \otimes \underline{X} \otimes \ldots \otimes \underline{X}) \\
&= \ell n(\underline{X}(1)\underline{X}(2) \ldots \underline{X}(N)) \\
&= \ell n\underline{X}(1) + \ell n\underline{X}(2) + \ldots + \ell n\underline{X}(N) \\
&= \underline{a}(1) + \underline{a}(2) + \ldots + \underline{a}(N) \\
&= \underline{A} \in \Lambda([GL(n)]^N).
\end{aligned}$$

The second line follows from the definition of $\underline{X}(k)$ (sec. I.4.3). The fourth line follows from $\ell n\underline{X}(k) \equiv \underline{a}(k)$ being well-defined; for write $\underline{X}(k) = \underline{E}(k) + \underline{X}'(k)$, then:

$$\ell n\underline{X}(k) = \ell n(\underline{E}(k) + \underline{X}'(k)) = \sum_q \frac{(-1)^{q-1}}{q} (\underline{E} \otimes \ldots \otimes \underline{X}' \otimes \ldots \otimes \underline{E})^q$$

$$= \underline{E} \otimes \underline{E} \otimes \ldots \otimes \ell n\underline{X}' \otimes \ldots \otimes \underline{E}$$

and, as is shown in the proof of lemma 1, $\ell n\underline{X}'$ exists; so it belongs to $(GL(n)) = AL(n,C)$. The last line follows from the theorem of sec. I.4.3.

Writing finally:

$$[\underline{X}]^N = e^{\ell n[\underline{X}]^N} = e^{\underline{A}} \equiv \sum_q \frac{1}{q!}\underline{A}^q, \qquad \underline{A} \in \Lambda([GL(n)]^N)$$

we see that $[\underline{X}]^N \in [N(\underline{E})]^N$ belongs to the enveloping algebra of $\Lambda([GL(n)]^N)$.

4 6. The interconnection between the tensor irreps of GL(n) and S_N

We have now laid the ground for the main theorem of this chapter. In fact everything regarding this theorem has already been proved and so the following result is not new, but a summary of the foregoing theory; it is for that purpose that it will be stated and proved explicitly.

Theorem.

Let $V_n{}^{\otimes N}$ be decomposed into S_N-irreducible subspaces $U_j^{[\lambda]}$ thus:

$$V_n{}^{\otimes N} = \sum_{(\lambda)} \oplus R^{(\lambda)}, \qquad R^{(\lambda)} = \sum_{j=1}^{n_{<\lambda>}} \oplus U_j^{[\lambda]} \tag{1}$$

Let the spaces $U_j^{[\lambda]}$, $j = 1, \ldots, n_{<\lambda>}$, span identical $f_{[\lambda]}$-dimensional irreps of S_N. Then the basis of $V_n{}^{\otimes N}$ adapted to the decomposition (1) is simultaneously adapted to the decomposition:

$$V_n{}^{\otimes N} = \sum_{(\lambda)} \oplus R^{(\lambda)}, \qquad R^{(\lambda)} = \sum_{i=1}^{f_{[\lambda]}} \oplus W_i^{<\lambda>} \tag{2}$$

where the spaces $W_i^{<\lambda>}$, $i = 1, \ldots f_{[\lambda]}$, span identical $n_{<\lambda>}$-dimensional irreps of GL(n), U(n) and SU(n).

Conversely, a decomposition of $V_n{}^{\otimes N}$ according to (2) induces a simultaneous decomposition according to (1).

Proof. The interdependence of the different representations is given by the following scheme:

$$S_N \overset{①}{\subset} CS_N \overset{②}{\longrightarrow} \boxed{AS_N \overset{③}{\leftrightarrow} AS_N^C \overset{④}{=} \mathfrak{C}([GL(n)]^N) \overset{⑤}{\supset} [GL(n)]^N} \overset{⑥}{\longleftarrow} GL(n) \overset{⑦}{\supset} U(n) \overset{⑧}{\supset} SU(n).$$

Inside the box the permutation group S_N (and its group algebra CS_N) standing on the left, and the Lie groups GL(n), U(n) and SU(n) standing on the right, are represented by linear operators on $V_n{}^{\otimes N}$.

Explanation of the different steps:

① S_N and its group algebra CS_N share their irreps (see e.g. ref. 9, 10 or 14).

② CS_N is represented on $V_n^{\otimes N}$ by AS_N (see sec. I.4.4).

③ Decomposition of $V_n^{\otimes N}$ under AS_N induces a simultaneous decomposition of $V_n^{\otimes N}$ under AS_N^C, the commutator algebra of AS_N. Also the converse of this statement holds (see sec. I.2.3).

④ AS_N^C is the enveloping algebra of $[GL(n)]^N$ (see sec. I.4.5).

⑤ The enveloping algebra of an operator group shares its irreps with the group (see sec. I.2.1).

⑥ GL(n) is represented on $V_n^{\otimes N}$ by $[GL(n)]^N$ (see sec. I.4.2).

⑦ GL(n) and U(n) share their irreps (see sec. I.3.3).

⑧ An irreducible tensor representation of U(n) stays irreducible under restriction to SU(n).

We show step 8:

The irreducible tensor representation $\underline{T}^{<\lambda>}(\underline{U})$, $\underline{U} \in U(n)$, is obtained from a similarity transformation on the matrix $[\underline{U}]^N \in [U(n)]^N$. Hence, the matrix elements of $\underline{T}^{<\lambda>}(\underline{U})$ are homogeneous polynomials of the order N in the matrix elements u_{ij} of $\underline{U}$. Now, any matrix $\underline{U} \in U(n)$ can be written as $\underline{U} = \alpha\underline{U}'$ with $\underline{U}' \in SU(n)$ (take $\alpha = (\det(\underline{U}))^{1/n}$), and so:

$$\underline{T}^{<\lambda>}(\underline{U}) = \underline{T}^{<\lambda>}(\alpha\underline{U}') = \alpha^N \underline{T}^{<\lambda>}(\underline{U}').$$

Applying the argument of ref. 9, p. 45 it follows that $\underline{T}^{<\lambda>}$ is an irreducible representation of SU(n) if and only if $\underline{T}^{<\lambda>}$ is an irreducible representation of U(n).

Note.

In this section only the irreps of the global groups have been discussed. Knowing that these irreps also constitute irreps of the corresponding Lie algebras (see sec. I.3.1), it follows that everything that has been stated in this section about the groups GL(n), U(n) and SU(n) also holds for the respective Lie algebras: $\Lambda(GL(n))$, $\Lambda(U(n))$ and $\Lambda(SU(n))$.

4.7. A note on the Casimir invariants of GL(n)

A well-known quantum mechanical problem is the determination of a complete set of commuting operators. The elements of such a set possess a set of common eigenvectors, which are uniquely determined (up to phase). Once a complete set of operators has been determined, the state labelling problem has been solved; i.e. every state is unambiguously characterized by a complete set

of eigenvalues. For instance the problem of labelling the atomic states arising from configurations of equivalent electrons led Racah in 1943 to the introduction of the seniority operator Q, which together with the angular momentum operators L^2 and L_z and spin operators S^2 and S_z yields an unambiguous labelling of states arising from d^n-configurations. In 1949 Racah was able to prove that Q, just like L^2, L_z, S^2 and S_z, is a Casimir invariant of a certain Lie group. This discovery initiated the search for, what is now called, canonical chains of groups, for the Casimir invariants, associated with such chains, form complete sets of commuting operators.

In this section we will show that the operators, representing the Casimir invariants of GL(n) on tensor space, belong to the center of AS_N. But before that, Casimir invariants will briefly be discussed in a more general context.

Definitions.

1. Any (abstract) Lie algebra Λ can be imbedded into an (abstract) associative algebra $\mathfrak{E}_U(\Lambda)$ in the following manner. Define a formal associative product on Λ and close Λ multiplicatively with respect to this product, thus constructing a semigroup of infinite order out of Λ. Close this semigroup linearly while defining $[a, b] = ab - ba$, $a, b \in \Lambda$, to relate the Lie product to the new associative product. The resulting infinite dimensional associative algebra $\mathfrak{E}_U(\Lambda)$ is the universal enveloping algebra of Λ.
2. The elements in the center of $\mathfrak{E}_U(\Lambda)$ are the Casimir invariants of Λ, i.e. they are elements $I \in \mathfrak{E}_U(\Lambda)$ such that $aI = Ia$, $\forall a \in \Lambda$; and hence I commutes with all elements of $\mathfrak{E}_U(\Lambda)$.

Notes.

1. See for a mathematically more satisfying definition of $\mathfrak{E}_U(\Lambda)$ ref. 29, p. 155.
2. A linear representation $\underline{D}$ of Λ extends to a representation of $\mathfrak{E}_U(\Lambda)$. $\mathfrak{E}_U(\Lambda)$ is represented by the enveloping algebra of $\underline{D}(\Lambda)$.
3. If $\underline{D}$ is irreducible then, by Schur's lemma, $\underline{D}(I) = \lambda_I \underline{E}$, for all Casimir invariants I of Λ.

Returning to the classical groups, we first note that GL(n) shares its Casimir invariants with U(n). This is obvious since the Lie algebra of

GL(n) is the complexification of the Lie algebra of U(n) (sec. I.3.2). So anything stated in the sequel about U(n) holds also for GL(n) (when considered as an n^2-dimensional complex Lie group).

Gelfand and later Biedenharn [30] have derived explicit expressions for a set of basic invariants of U(n). ('Basic' means here that any Casimir invariant of U(n) can be expressed as an analytic function of these invariants, which themselves are functionally independent). It can be proved [30] that U(n) has n and that SU(n) has n-1 basic invariants. Denoting the operators representing the invariants $\{I_i^{(n)}\}$ of U(n) on $V_n{}^{\otimes N}$ by $\{\hat{I}_i^{(n)}\}$ it follows from Schur's lemma that the $(f_{[\lambda]} \times n_{<\lambda>})$-dimensional spaces $R^{(\lambda)}$, defined in sec. I.4.6 as the direct sum of all the spaces carrying the irrep $[\lambda]$ of S_N, are eigenspaces of the n operators $\{\hat{I}_i^{(n)}\}$. Biedenharn proved that the n corresponding eigenvalues give a unique designation of the irreps of U(n). These eigenvalues have been computed explicitly by Louck [31]; they are given as rather complicated expressions in terms of partial hooks [31]. (Recently Hudson [32] derived another system of basic invariants for U(n), of which the eigenvalues are simple polynomials in the components of the highest weight of the irrep).

Of course the n operators $\{\hat{I}_i^{(n)}\}$ representing the Casimir invariants of U(n) do not constitute a complete set of operators. However, the $\frac{1}{2}n(n+1)$ invariants belonging to the canonical chain:

$$U(n) \supset U(n-1) \supset \ldots \supset U(1)$$

form a complete set of operators (we come back to this point in sec. I.5.6).

The operator S^2, representing the (only) Casimir invariant of SU(2) on N-electron spin space, can be expressed as a linear combination of permutation operators. This can be derived from the Dirac identity [33, p. 222]. The following theorem can be considered as an alternative derivation and a generalization of this result.

Theorem.

The operators on $V_n{}^{\otimes N}$ that represent the Casimir invariants of GL(n) and U(n) belong to the center $A_c S_N$ of AS_N.

Proof. An invariant I belongs by definition to the center of the universal enveloping algebra of the Lie algebra of GL(n). Hence $\hat{I}$, representing I on $V_n{}^{\otimes N}$, belongs to the center of the enveloping algebra of $A([GL(n)]^N)$. In sec. I.4.5 it has been proved that this enveloping algebra coincides with AS_N^c, the commutator algebra of AS_N. The center of an algebra being

the intersection of the algebra and its commutator algebra, we get:

$$\hat{I} \in AS_N^C \cap (AS_N^C)^C .$$

From sec. I.2.3 we know that $(AS_N^C)^C = AS_N$ (AS_N is semi-simple, as it represents the finite group algebra CS_N), and therefore:

$$\hat{I} \in (AS_N \cap AS_N^C) \equiv A_C S_N .$$

Notes.

1. The map of the Casimir invariants of U(n) is usually into $A_C S_N$. So, in general, the set operators $\{\hat{I}_i^{(n)}\}$ spans a proper subspace of $A_C S_N$.
2. Two possible bases of $A_C S_N$ are the set of class sum operators and the set of character projectors of S_N (see e.g. ref. 10, theorems 5.3, 5.4). The operators $\{\hat{I}_i^{(n)}\}$ can be expressed in terms of either of these two sets.

Example.

The two spin orbitals α and β span the 2-dimensional spin space V_2, which carries an irrep of $\Lambda(SU(2))$ and its complexification $\Lambda_C(SU(2))$. To clarify the theory of this chapter we discuss some points regarding the decomposition of $V_2 \otimes^4$, the 4-electron spin space.

From standard angular momentum theory [17, sec. 3.1] it is known how to decompose $V_2 \otimes^4$ under $\Lambda_C(SU(2))$. Briefly, the procedure is the following:

(i) Choose the vector with highest eigenvalue of S_z (there is only one such vector in spin space).

(ii) Apply the step-down operator:

$$S_- = \sum_k s_-(k), \qquad s_-(k) = s_x(k) - i s_y(k)$$

repeatedly onto this vector until the zero vector is generated. Because $S_- \in \Lambda_C([SU(2)]^N)$ (see sec. I.4.3), the space V_S so generated carries an irrep of $\Lambda_C(SU(2))$ and of $\Lambda(SU(2))$. SU(2) being the universal covering group (sec. I.3.1 and I.3.4) of all the groups with this same Lie algebra, V_S carries also an irrep of SU(2).

(iii) Choose a vector with the highest but one eigenvalue of S_z, orthogonalize to V_S, apply the step-down operator, etc.

Proceeding in this manner, one finds that $V_2 \otimes^4$ decomposes into SU(2)-irreducible spaces thus:

$$V_2 \otimes^4 = V'_{S=2} \oplus V'_{S=1} \oplus V''_{S=1} \oplus V'''_{S=1} \oplus V'_{S=0} \oplus V''_{S=0}$$

with corresponding basis:

$V'_{S=2}$:	$\|2,2,1\rangle$	$\|2,1,1\rangle$	$\|2,0,1\rangle$	$\|2,-1,1\rangle$	$\|2,-2,1\rangle$
$V'_{S=1}$:		$\|1,1,1\rangle$	$\|1,0,1\rangle$	$\|1,-1,1\rangle$	
$V''_{S=1}$:		$\|1,1,2\rangle$	$\|1,0,2\rangle$	$\|1,-1,2\rangle$	
$V'''_{S=1}$:		$\|1,1,3\rangle$	$\|1,0,3\rangle$	$\|1,-1,3\rangle$	
$V'_{S=0}$:			$\|0,0,1\rangle$		
$V''_{S=0}$:			$\|0,0,2\rangle$		

where $|S,M_S,\iota\rangle$ denotes an eigenvector of S^2 with quantum number S, and of S_z with quantum number M_S; ι labels the remaining multiplicity. Applying now the theorem of sec. I.4.6 we see that each of the vectors $|2,M_S,1\rangle$, $M_S = -2, \ldots, +2$, spans the same 1-dimensional irrep of S_4; the 3-dimensional sets: $\{|1,1,\iota\rangle\}$, $\{|1,0,\iota\rangle\}$ and $\{|1,-1,\iota\rangle\}$ ($\iota = 1,2,3$) span the same 3-dimensional irrep of S_4; and $\{|0,0,\iota\rangle\}$ ($\iota = 1,2$) spans a 2-dimensional irrep of S_4.

The Casimir operator S^2 can be expressed in terms of the character projectors (sec. I.2.2) of S_4:

$$S^2 = \sum_{[\lambda]} c_\lambda \, e^{[\lambda]} .$$

Applying S^2 in this form to a projected vector $e^{[\mu]}(v)$, $v \in V_2 \otimes^4$, further using $e^{[\lambda]} e^{[\mu]} = \delta^{\lambda,\mu} e^{[\lambda]}$ together with the fact that $e^{[\mu]}(v)$ is an eigenvector of S^2, it follows that the expansion coefficients c_λ are eigenvalues of S^2. In ch. I.6 the irreps of S_N will be brought into correspondence to the eigenvalues $S(S+1)$ of S^2, anticipating that result we get

$$S^2 = 2(2+1)e^{[4]} + 1(1+1)e^{[3,1]} + 0(0+1)e^{[2,2]} .$$

One can also rewrite this expression in terms of class sum operators. After some manipulation one arrives at:

$$S^2 = C^{(2,1^2)} \equiv (12) + (13) + (14) + (23) + (24) + (34).$$

This result could also have been obtained directly by applying the Dirac identity [33].

CHAPTER I.5 THE DECOMPOSITION OF TENSOR SPACE

In the foregoing chapter it has been shown that the decomposition of $V_n^{\otimes N}$ under S_N goes hand in hand with the decomposition under the Lie groups GL(n), SU(n), U(n) and their respective Lie algebras. In this chapter it will be discussed how bases of subspaces of $V_n^{\otimes N}$, carrying irreps of these groups, can be constructed in practice.

In principle this purpose may be achieved in one of the following three ways:

(i) Use Lie algebraic techniques to decompose $V_n^{\otimes N}$ under $\Lambda(U(n))$. This decomposition may then be followed by a computation of the matrix elements of the generators of U(n) over the irreducible bases thus constructed. Since every linear operator on $V_n^{\otimes N}$ can be expressed in terms of the generators of U(n), the matrix representation of every (bisymmetric) linear operator can then be calculated. The full power of this approach has yet to be explored by quantum chemists, although recently discussions of this method, applied to N-electron systems, have been given by Paldus [34] and Gouyet et al. [35].

(ii) Use an approach from the point of view of the global group U(n), generalizing the necessary results of finite groups to this compact group. This method requires very sophisticated mathematics.

(iii) Use the representation theory of S_N, developed by Young, Littlewood and others.

The last method is probably the most convenient of the three, and is in any case by far the best known among theoretical chemists. We will also follow this line of approach. To that end some of the necessary theory regarding the decomposition of the group algebra CS_N of S_N is introduced in the first three sections. We rely heavily on the book of Rutherford [36] for that, although not all the results to be presented can be found in that reference. Two different basis sets corresponding to a complete decomposition of the group algebra of CS_N will be discussed.

The first consists of Young units of the NP-type, sometimes called "structure projectors" [37], leading to the "Weyl-Rumer basis", the spin-free equivalent of "spin-bonded" functions [38-40]. This basis has been used in the computer calculations presented in the second half of this thesis.

The second basis of CS_N to be discussed consists of Young-Yamanouchi units, which carry orthogonal representations of S_N. These units give rise to a basis of $V_n^{\otimes N}$ known as the Gelfand basis.

The definitions of Young diagrams, (standard) Young tableaux etc., especially in connection with their use for labelling the irreps of S_N, are presupposed; a good reference for this is an article by Coleman [41]. We just note that the term Young diagram applies to a frame with empty boxes, while a Young tableau consists of N boxes, filled with the numbers 1 to N.

5.1. Young units

Definitions.

1. All possible permutations moving numbers along the rows of a Young tableau $T_r^{[\lambda]}$ form a group $R_r^{[\lambda]}$: the row group belonging to the tableau $T_r^{[\lambda]}$. The elements of $R_r^{[\lambda]}$ are called horizontal permutations.
2. All possible permutations moving numbers along the columns of a Young tableau $T_r^{[\lambda]}$ (vertical permutations) form a group $C_r^{[\lambda]}$: the column group belonging to $T_r^{[\lambda]}$.
3. The element
$$P_r^{[\lambda]} = \Sigma\, \pi,$$
where the sum runs over all $\pi \in R_r^{[\lambda]}$, is the row symmetrizer belonging to $T_r^{[\lambda]}$.
4. The element
$$N_r^{[\lambda]} = \Sigma\, \zeta_\nu\, \nu,$$
where ζ_ν is the parity of ν and the sum runs over all $\nu \in C_r^{[\lambda]}$, is the column antisymmetrizer belonging to $T_r^{[\lambda]}$.

Example.

$T_r^{[2,2]} = \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 4 & 2 \\ \hline \end{array}$ (a non standard Young tableau).

$R_r^{[2,2]} = \{(1), (13), (24), (13)(24)\}$, $P_r^{[2,2]} = (1) + (13) + (24) + (13)(24)$

$C_r^{[2,2]} = \{(1), (34), (12), (12)(34)\}$, $N_r^{[2,2]} = (1) - (34) - (12) + (12)(34)$.

Definition.

Let $T_s^{[\lambda]}$ and $T_t^{[\lambda]}$ be Young tableaux with shape $[\lambda]$, then we define the permutation $\sigma_{ts}^{[\lambda]}$ by:

$$\sigma_{ts}^{[\lambda]}\, T_s^{[\lambda]} = T_t^{[\lambda]}, \qquad \sigma_{ts}^{[\lambda]} \in S_N.$$

Example.

$$T_s^{[\lambda]} = \begin{array}{|c|c|c|}\hline 1 & 3 & 2 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}, \quad T_t^{[\lambda]} = \begin{array}{|c|c|c|}\hline 2 & 5 & 3 \\ \hline 4 & 1 \\ \cline{1-2}\end{array} \Rightarrow \sigma_{ts}^{[3,2]} = (1235).$$

(Note the permutation convention; $\sigma_{ts}^{[\lambda]}$ operates on the numbers. Because a Young tableau contains by definition N different numbers no confusion is possible).

Definition.

Let $T_s^{[\lambda]}$ and $T_t^{[\lambda]}$ be Young tableaux with shape $[\lambda]$, then:

$$Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]}\, N_s^{[\lambda]}\, P_s^{[\lambda]}$$

is a Young unit of the NP-type.

Note.

The definition:

$$Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]}\, P_s^{[\lambda]}\, N_s^{[\lambda]}$$

gives a Young unit of the PN-type.

Lemma.

$$Y_{ts}^{[\lambda]} = \sigma_{ts}^{[\lambda]}\, N_s^{[\lambda]}\, P_s^{[\lambda]} = N_t^{[\lambda]}\, \sigma_{ts}^{[\lambda]}\, P_s^{[\lambda]} = N_t^{[\lambda]}\, P_t^{[\lambda]}\, \sigma_{ts}^{[\lambda]}.$$

Proof. Ref. 36, p. 16.

From here on the subscripts of the Young units will be running over standard tableaux only, unless stated differently. Recall in this connection that $f_{[\lambda]}$ standard tableaux of shape $[\lambda]$ can be constructed, where $f_{[\lambda]}$ is the dimension of the irrep $[\lambda]$ of S_N.

Definition.

Since the Young units belong to CS_N, they can be expressed as a linear combination of permutations:

$$Y_{rs}^{[\lambda]} = \sum_{P \in S_N} U^{[\lambda]}(P)_{rs}\, P.$$

Letting r and s run over standard tableaux and keeping P fixed, the expansion coefficients $U^{[\lambda]}(P)_{rs}$ form an $f_{[\lambda]} \times f_{[\lambda]}$ matrix $\underline{U}^{[\lambda]}(P)$.

Notes.

1. The matrix $\underline{U}^{[\lambda]}(P)$ is not a matrix representation of P.
2. From its definition the matrix $\underline{U}^{[\lambda]}(P)$ can be calculated by writing out all $f_{[\lambda]}^2$ units. A shorter route is described by means of the following example.

Example.

Tableaux $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$ and $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$ give rise to the Young units:

$Y_{11} = (1) + (12) - (13) - (123)$ $\quad Y_{12} = (23) - (132) + (123) - (12)$

$Y_{21} = (23) + (132) - (123) - (13)$ $\quad Y_{22} = (1) - (12) + (13) - (132)$

and so for instance:

$$\underline{U}^{[2,1]}(123) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \quad \underline{U}^{[2,1]}(132) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$$

A direct way of computing the transposed matrix $\underline{U}^{[\lambda]}(P)^T$ for a given permutation P is the following:

Construct an $f_{[\lambda]} \times f_{[\lambda]}$ table in which the columns are labelled by standard tableaux $T_r^{[\lambda]}$, $r = 1, 2, \ldots, f_{[\lambda]}$ and the rows by the permuted tableaux $P(T_r^{[\lambda]})$, $r = 1, \ldots, f_{[\lambda]}$, thus:

(123):

	$\begin{array}{\|c\|c\|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$
$\begin{array}{\|c\|c\|}\hline 2&3\\\hline 1\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 3&2\\\hline 1\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 2&3\\\hline 1\\\cline{1-1}\end{array}$
$\begin{array}{\|c\|c\|}\hline 2&1\\\hline 3\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$	0

(132):

	$\begin{array}{\|c\|c\|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$
$\begin{array}{\|c\|c\|}\hline 3&1\\\hline 2\\\cline{1-1}\end{array}$	0	$\begin{array}{\|c\|c\|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$
$\begin{array}{\|c\|c\|}\hline 3&2\\\hline 1\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 3&2\\\hline 1\\\cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|}\hline 2&3\\\hline 1\\\cline{1-1}\end{array}$

The tableaux contained inside the tables are obtained by application of certain horizontal permutations to the tableaux labelling the rows. These horizontal permutations are chosen such that after permutation the numbers appear in the same column as in the tableau at the top. If no such horizontal permutation exists, that is, if two numbers would have to appear in the same box, we write down a zero in the table. The tableaux thus obtained can now be transformed to the standard tableaux at the top via a vertical permutation ν with parity ζ_ν. The matrix $\underline{U}^{[\lambda]}(P)^T$ follows finally by replacing the tableaux in the table by this parity, so:

$$\underline{U}(123)^T = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \qquad \underline{U}(132)^T = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

The proof of this construction is given by the NP-equivalence of Rutherford's theorem 7 [36].

The Young units that belong to the same diagram $[\lambda]$ are not necessarily orthogonal. Still, by virtue of Von Neumann's theorem [36, theorem 8], they satisfy a rather simple multiplication rule. This is the content of the following theorem.

Theorem.

$$Y^{[\lambda]}_{ru}\, Y^{[\mu]}_{vs} = \frac{N!}{f_{[\lambda]}}\, \xi^{[\lambda]}_{vu}\, \delta^{\lambda,\mu}\, Y^{[\lambda]}_{rs}\,.$$

Here $\xi^{[\lambda]}_{vu}$ is the coefficient of the identity (1) in $Y^{[\lambda]}_{vu}$.

Proof. Translate Rutherford's theorem 11 [36] into the NP-definition of a Young unit.

Note.

The numbers $\xi^{[\lambda]}_{vu}$, $v, u = 1, \ldots, f_{[\lambda]}$, constitute an $f_{[\lambda]} \times f_{[\lambda]}$ matrix $\underline{X}^{[\lambda]}$, the "NP-structure matrix", which will play an important rôle in this and the following chapter. It is easy to show from its definition that the matrix elements of $\underline{X}^{[\lambda]}$ are given by the following rules. Assume to that end that the standard tableaux are in dictionary order (defined in ref. 8, ch. IV, §4) and suppress $[\lambda]$ for the moment. Then

$u = v$ $\quad \xi_{vv} = 1$

$u > v$ $\quad \xi_{uv} = 0$

$u < v$ $\quad \xi_{uv} = \tau_{ut}$, if T_v does not have a pair of numbers in one row that T_u has in one column.

$\xi_{uv} = 0$, if T_v has a pair of numbers in one row that T_u has in one column.

The quantity τ_{ut}, which is the parity of the permutation σ_{ut}, arises from the following consideration. If $\xi_{uv} \neq 0$ then σ_{uv} can be written as [36, theorem 5]:

$$\sigma_{uv} = \nu_u \pi_v, \qquad \nu_u \in C_u,\ \pi_v \in R_v.$$

Defining now the tableau T_t by:

$$T_t = \pi_v T_v$$

we see that $\sigma_{ut} = \nu_u$, for: $T_u = \sigma_{uv} T_v = \nu_u \pi_v T_v = \nu_u T_t$, and so the parity τ_{ut} of σ_{ut} is equal to the parity of ν_u.

A mechanical way to obtain $\underline{X}^{[\lambda]}$ is the same as for $\underline{U}^{[\lambda]}(P)$ if we take $P = (1)$. Label the rows of an $f_{[\lambda]} \times f_{[\lambda]}$ array with the standard tableaux T_v, $v = 1, \ldots, f_{[\lambda]}$ and the columns with the standard tableaux T_u, $u = 1, \ldots, f_{[\lambda]}$. Apply the horizontal permutation π_v to T_v, with $\pi_v T_v = T_t$, so that the digits in T_t appear, if possible, in the same column as in T_u. Replace the tableau T_t thus obtained by $\zeta_{\nu_u} = \tau_{ut}$. Transpose the result.

The NP-structure matrix is very easily determined, as can be seen in the following example.

Example.

T_v \ T_u	1 2 3 4 5	1 2 3 5 4	1 3 2 4 5	1 3 2 5 4	1 4 2 5 3
1 2 3 4 5	1 2 3 4 5				
1 2 3 5 4	0	1 2 3 5 4			
1 3 2 4 5	0	0	1 3 2 4 5		
1 3 2 5 4	0	0	0	1 3 2 5 4	
1 4 2 5 3	1 4 5 2 3	0	0	0	1 4 2 5 3

$$\Rightarrow \underline{X}^{[2^2,1]} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Notes.

1. The matrix $\underline{X}^{[\lambda]}$ is non-singular, for because it is upper triangular and has +1 along the diagonal its determinant is +1.
2. The inverse matrix $(\underline{X}^{[\lambda]})^{-1}$ is also upper triangular.
3. The elements of this inverse matrix will be denoted by $\eta_{rs}^{[\lambda]}$. They also will be used frequently in the sequel.

As will be stated in the following theorem, the Young units form a basis of CS_N corresponding to a simultaneous decomposition of this semi-simple algebra into minimal left ideals, right ideals and simple invariant subalgebras. So, except for the fact that they lack orthogonality, Young units give a Weddeburn decomposition of CS_N (see sec. I.2.2).

Theorem.

(i) The "diagonal" elements:

$$\frac{f_{[\lambda]}}{N!} Y^{[\lambda]}_{rr}, \qquad \text{all } [\lambda],\ r = 1, \ldots, f_{[\lambda]}$$

are primitive idempotents.

(ii) The set:

$$\{Y^{[\lambda]}_{rs} \mid s = 1, \ldots, f_{[\lambda]},\ [\lambda] \text{ and } r \text{ fixed}\}$$

spans a minimal right ideal in CS_N.

(iii) The set:

$$\{Y^{[\lambda]}_{rs} \mid r = 1, \ldots, f_{[\lambda]},\ [\lambda] \text{ and } s \text{ fixed}\}$$

spans a minimal left ideal in CS_N.

(iv) The Young units afford a resolution of the identity:

$$(1) = \sum_{[\lambda]} \frac{f_{[\lambda]}}{N!} \sum_{r=1}^{f_{[\lambda]}} \sum_{t=1}^{f_{[\lambda]}} \eta^{[\lambda]}_{tr}\, Y^{[\lambda]}_{rt}.$$

(v) The set:

$$\{Y^{[\lambda]}_{rs} \mid \text{all } [\lambda],\ r, s = 1, \ldots, f_{[\lambda]}\}$$

forms a complete linearly independent basis of CS_N.

Proof.

(i) Translate Boerner's theorem 3.1 [8, chapter IV] into the NP-definition of a Young unit.

(ii) and (iii) Apply the manner of proof used at p. 102 of ref. 42.

(iv) Ref. 36, § 29. Note in this respect that

$$g^{[\lambda]}_{rr} \equiv \sum_t \eta^{[\lambda]}_{rt}\, Y^{[\lambda]}_{tr}$$

is a natural unit.

(v) Apply (iv) together with theorem 4.6 in chapter IV of ref. 8.

Definition.

Because of the theorem just stated the Young units span irreducible

representations of CS_N, with the elements of CS_N acting from the right:

$$Y^{[\lambda]}_{rs}\, Q = \sum_{t=1}^{f^{[\lambda]}} {}_R D^{[\lambda]}(Q)_{ts}\, Y^{[\lambda]}_{rt}, \qquad Q \in S_N$$

and also from the left:

$$Q\, Y^{[\lambda]}_{rs} = \sum_{t=1}^{f^{[\lambda]}} {}_L D^{[\lambda]}(Q)_{tr}\, Y^{[\lambda]}_{ts}, \qquad Q \in S_N.$$

The irreducible matrix ${}_R\underline{D}(Q)$ defined this way can easily be computed, using the following theorem.

Theorem.

$${}_R\underline{D}^{[\lambda]}(Q) = (\underline{X}^{[\lambda]})^{-1}\, \underline{U}^{[\lambda]}(Q^{-1}),$$

where $\underline{X}^{[\lambda]}$ is the NP-structure matrix previously defined, and $\underline{U}^{[\lambda]}(Q^{-1})$ contains the components of the Young units along Q^{-1}.

Proof. Drop $[\lambda]$.

$$\sum_{P \in S_N} U(P)_{rs}\, PQ = Y_{rs}\, Q = \sum_{t=1}^{f} {}_R D(Q)_{ts}\, Y_{rt} = \sum_t \sum_P U(P)_{rt}\, {}_R D(Q)_{ts}\, P.$$

Substitute $R = PQ$ and $P = RQ^{-1}$ in the leftmost expression and $P = R$ in the rightmost expression:

$$\sum_{R \in S_N} U(RQ^{-1})_{rs}\, R = \sum_{R \in S_N} \{ \sum_{t=1}^{f} U(R)_{rt}\, {}_R D(Q)_{ts} \} R.$$

Compare coefficients of $R = (1)$:

$$U(Q^{-1})_{rs} = \sum_t U(1)_{rt}\, {}_R D(Q)_{ts}.$$

By definition $U(1)_{rt}$ is the coefficient of (1) in the unit Y_{rt}, also by definition this coefficient is ξ_{rt} and so:

$$\underline{U}(Q^{-1}) = \underline{X}\, {}_R\underline{D}(Q).$$

Note.

In the same way one proves for the left-irrep ${}_L\underline{D}(Q)$

$${}_L\underline{D}(Q)^T = \underline{U}(Q^{-1})\ (\underline{X})^{-1}.$$

5.2. Young orthogonal units

Since it is often advantageous to work with an orthogonal basis of CS_N which moreover carries orthogonal irreps, Young defined an alternative basis for CS_N, satisfying these two conditions. (Note that an orthogonal basis of a group algebra does not necessarily carry orthogonal representations!)

Definition.

The following recursive expression for an orthogonal Young unit $o^{[\lambda]}_{pq}$ is due to Young:

$$o^{[\lambda]}_{pq} = \frac{f^{[\lambda]}}{N!} A^{[\lambda]}_{pq} \; o^{[\lambda(N-1)]}_{pp} \; Y^{[\lambda(N)]}_{pq} \; o^{[\lambda'(N-1)]}_{qq} .$$

Here:

$[\lambda(N)] \equiv [\lambda]$ is an irrep of S_N,

$[\lambda(N-1)]$ is the irrep of S_{N-1} with a diagram obtained from $T^{[\lambda(N)]}_{p}$ by removal of the digit N,

$[\lambda'(N-1)]$ is the irrep of S_{N-1} obtained by removal of the digit N from $T^{[\lambda(N)]}_{q}$,

$Y^{[\lambda(N)]}_{pq}$ is a Young unit,

$A^{[\lambda]}_{pq}$ is a normalizing constant (see Rutherford §26 [36] for the computation).
Note: $A^{[\lambda]}_{pq}$ in the NP-definition of $Y^{[\lambda(N)]}_{pq}$ is equal to $(A^{[\lambda]}_{pq})^{-1}$ in the PN-definition.

Example.

$T^{[\lambda(4)]}_{1} = \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4\\\cline{1-1}\end{array}$, $T^{[\lambda(3)]}_{1} = \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}$, $T^{[\lambda(2)]}_{1} = \begin{array}{|c|c|}\hline 1&2\\\hline\end{array}$, $T^{[\lambda(1)]}_{1} = \begin{array}{|c|}\hline 1\\\hline\end{array}$

$T^{[\lambda(4)]}_{2} = \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3\\\cline{1-1}\end{array}$, $T^{[\lambda'(3)]}_{1} = \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$, $T^{[\lambda'(2)]}_{1} = \begin{array}{|c|c|}\hline 1&2\\\hline\end{array}$, $T^{[\lambda'(1)]}_{1} = \begin{array}{|c|}\hline 1\\\hline\end{array}$

So:

$o^{[\lambda(1)]}_{11} \equiv o^{[1]}_{11} = (1)$;

$o^{[\lambda(2)]}_{11} \equiv o^{[2]}_{11} = 1/2[(1) + (12)]$;

$o^{[\lambda(3)]}_{11} \equiv o^{[3]}_{11} = 1/6[(1) + (12) + (13) + (23) + (123) + (132)]$;

$o^{[\lambda'(3)]}_{11} \equiv o^{[2,1]}_{11} = 1/3[(1) + (12) - 1/2\{(13) + (23) + (123) + (132)\}]$;

$o^{[\lambda(4)]}_{12} \equiv o^{[3,1]}_{12} = 3/24 \cdot 3/4\sqrt{2} \; o^{[3]}_{11} Y^{[3,1]}_{12} o^{[2,1]}_{11}$, where $3/4\sqrt{2} = A^{[3,1]}_{12}$,

and so:

$$o_{12}^{[\lambda(4)]} = -1/3\sqrt{2}\{(14) + (24) - 2\cdot(34) + (124) - 2\cdot(134) + (142) + (143) - 2\cdot(234) + (243) - 2\cdot(1234) + (1243) + (1324) - 2\cdot(1342) + (1423) + (1432) - 2\cdot(12)(34) + (13)(24) + (14)(23)\}.$$

The orthogonal Young units belonging to standard tableaux form a basis of CS_N corresponding to a Wedderburn decomposition of this group algebra. The following expression, here written in terms of orthogonal Young units, can be shown to hold in general for orthogonal units of semi-simple algebras (e.g. [8], p. 66):

$$o_{pq}^{[\lambda]} = \frac{f^{[\lambda]}}{N!} \sum_{P \in S_N} D^{[\lambda]}(P^{-1})_{qp}\, P.$$

Here $\underline{D}^{[\lambda]}$ is the matrix representation of S_N carried by the orthogonal units. The orthogonal Young units have, by virtue of the specific manner in which they were defined, the property that $\underline{D}^{[\lambda]}(P^{-1})$ is a real orthogonal matrix [36, § 27].

Denoting the trace of $\underline{D}^{[\lambda]}(P^{-1})$ by $\chi^{[\lambda]}(P^{-1})$ it follows that

$$o^{[\lambda]} \equiv \sum_{q=1}^{f^{[\lambda]}} o_{qq}^{[\lambda]} = \frac{f^{[\lambda]}}{N!} \sum_{P \in S_N} \chi^{[\lambda]}(P^{-1})\, P.$$

So: $o_{pq}^{[\lambda]}$ is a Wigner operator and $o^{[\lambda]}$ is a character projector.

Theorem.

The diagonal units satisfy the recursion relation:

$$o_{qq}^{[\lambda(N)]} = o^{[\lambda(N)]}\, o_{qq}^{[\lambda(N-1)]},$$

where $o_{qq}^{[\lambda(N-1)]}$ belongs to the standard tableau obtained from $T_q^{[\lambda(N)]}$ by removing the number N.

Proof. (The manner of proof is inspired by chapter VII of ref. 11).

First we note that the element

$$o^{[\lambda(N)]}\, o_{qq}^{[\lambda(N-1)]}$$

is a primitive idempotent.

The element is idempotent since the factors commute and both are idempotent. (Recall here that $o^{[\lambda(N)]}$ belongs to the center of CS_N).

The element is primitive because of the following reason: The simple algebra obtained by the action of $o^{[\lambda(N-1)]}$ onto CS_{N-1}:

$$A^{[\lambda(N-1)]} \equiv o^{[\lambda(N-1)]} \quad (CS_{N-1})$$

occurs just once in the simple algebra:

$$A^{[\lambda(N)]} \equiv o^{[\lambda(N)]} \quad (CS_N),$$

because the irrep $[\lambda(N-1)]$ occurs just once in the restriction of $[\lambda(N)]$ to S_{N-1}. Of course $A^{[\lambda(N-1)]}$ possesses only one q-th diagonal unit, and so the element $o^{[\lambda(N)]}\, o^{[\lambda(N-1)]}_{qq}$ belongs to a one-dimensional subspace of CS_N, from which follows that it is primitive.

We can now write:

$$o^{[\lambda(N)]}_{qq} \equiv o^{[\lambda(N-1)]}_{qq}\, Y^{[\lambda(N)]}_{qq}\, o^{[\lambda(N-1)]}_{qq} = o^{[\lambda(N-1)]}_{qq}\, o^{[\lambda(N)]}\, Y^{[\lambda(N)]}_{qq}\, o^{[\lambda(N)]}\, o^{[\lambda(N-1)]}_{qq}$$

$$= c\; o^{[\lambda(N)]}\, o^{[\lambda(N-1)]}_{qq}, \qquad c \in C.$$

The first equal sign follows from the facts that $o^{[\lambda(N)]}$ is a unit of $A^{[\lambda(N)]}$ and that $Y^{[\lambda(N)]}_{qq}$ belongs to this algebra. The second equal sign follows (via theorem 3.9, chapter III of ref. 8) from the fact that:

$$o^{[\lambda(N)]}\, o^{[\lambda(N-1)]}_{qq}$$

is a primitive idempotent. The constant c is equal to 1 since both sides in the relation to be proved are idempotent.

Notes.

1. Continuing the recursion it follows that:

$$o^{[\lambda(N)]}_{qq} = o^{[\lambda(N)]}\, o^{[\lambda(N-1)]}\, o^{[\lambda(N-2)]} \ldots o^{[1]},$$

   so there exists a 1-1 correspondence between the index q and the set of irreps $\{[\lambda(N)], [\lambda(N-1)], \ldots, [1]\}$, which forms the genealogy of the q-th standard Young tableau with shape $[\lambda(N)]$.

2. The recursive expression in the foregoing note has been given earlier by Matsen [25], who, after having defined the diagonal units in this manner, proves that they are primitive idempotents. In this manner I cannot prove, however, that the diagonal units are the same as the orthogonal units.

5.3. Yamanouchi units

Apparently not aware of the work of Young, Yamanouchi [43] derived a set of orthogonal irreps of S_N by using a sort of vector coupling technique. In

this section it will be proved, by purely algebraic means, that Young's orthogonal and Yamanouchi's representation are identical.

Definition.

A Yamanouchi unit $W^{[\lambda]}_{rs}$ is defined by:

$$W^{[\lambda]}_{rs} = \frac{f^{[\lambda]}}{N!} \sum_{P \in S_N} O^{[\lambda]}(P)_{rs} P,$$

where the matrix $\underline{O}^{[\lambda]}(P)$ is by Yamanouchi's construction adapted to the sequence

$$S_N \supset S_{N-1} \supset \ldots \supset S_1.$$

Hence Yamanouchi units are orthogonal.

To prove that $W^{[\lambda]}_{rs}$ is in fact a Young orthogonal unit, we first note that the diagonal units can be written thus:

$$W^{[\lambda(N)]}_{rr} = \sum_{q=1}^{f^{[\lambda(N-1)]}} \tilde{W}^{[\lambda(N)]}_{rq} W^{[\lambda(N-1)]}_{qr}$$

(where $\tilde{W}^{[\lambda(N)]}_{rq}$ is a "reduced" Yamanouchi unit, see the appendix of ref. 44), and therefore:

$$W^{[\lambda(N)]} W^{[\lambda(N-1)]}_{tt} = \left(\sum_{r=1}^{f^{[\lambda]}} W^{[\lambda(N)]}_{rr} \right) W^{[\lambda(N-1)]}_{tt} = \sum_r \sum_q \tilde{W}^{[\lambda(N)]}_{rq} W^{[\lambda(N-1)]}_{qr} W^{[\lambda(N-1)]}_{tt}$$

$$= W^{[\lambda(N)]}_{tt}.$$

From which follows that the diagonal Yamanouchi units satisfy the same recursion relation as the diagonal Young orthogonal units; hence the two sets of diagonal units are identical.

We now show that the same holds for the off-diagonal units. Write to that end, suppressing $[\lambda]$.

$$W_{ij} = \sum_{k,\ell} c_{k\ell;ij} O_{k\ell}.$$

Multiply on the left by $W_{ii} = O_{ii}$ and on the right by $W_{jj} = O_{jj}$:

$$W_{ii} W_{ij} W_{jj} = W_{ij} = \sum_{k,\ell} c_{k\ell;ij} O_{ii} O_{k\ell} O_{jj} = c_{ij;ij} O_{ij}.$$

Both sets of units carry real orthogonal representations:

$$\bar{W}_{ij} = W_{ji} \Rightarrow (c_{ij;ij} O_{ij})^{\perp} = c_{ji;ji} O_{ji} \Rightarrow c_{ij;ij} = c_{ji;ji} \equiv c_{ij}.$$

Further.

$$o_{ii} = W_{ii} = W_{ij}W_{ji} = c^2_{ij}o_{ij}o_{ji} = c^2_{ij}o_{ii} \Rightarrow c_{ij} = \pm 1.$$

Now, it has already been pointed out by Yamanouchi himself [43, p. 422] that the requirement that $[\lambda]$ must be adapted to

$$S_N \supset S_{N-1} \supset \ldots \supset S_1,$$

only determines $[\lambda]$ up to a similarity transformation with a diagonal matrix with elements ± 1 along the diagonal. This implies that the sign of the off-diagonal units cannot be fixed unambiguously. However, the choice: $c_{ij} = +1$ yields the convention chosen by Yamanouchi. (I am indebted to Dr.D.J. Klein for drawing my attention to this freedom in sign).

Concluding we have found that Young orthogonal units and Yamanouchi units are identical. Henceforth they will be called Young-Yamanouchi units, or briefly Y-Y units.

5.4. Graphical representation of projected tensors

In the preceding three sections two possible complete decompositions of CS_N have been given. Explicit expressions for the corresponding basis elements, Young units and Y-Y units, have been derived. From the point of view of the mathematician the problem of decomposing tensor space under GL(n), U(n) and SU(n) has now been solved completely, as the irreducible GL(n)-carrier space $W_1^{<\lambda>}$, defined in sec. I.4.6, is simply given by:

$$W_1^{<\lambda>} = e_{11}^{[\lambda]} (V_n \otimes^N),$$

where $e_{11}^{[\lambda]}$ stands for either $Y_{11}^{[\lambda]}$ (a Young unit) or $W_{11}^{[\lambda]}$ (a Y-Y unit).

However, projection is usually a very tedious procedure, because many times a zero vector or a vector linearly dependent on the vectors already projected will be found. So, in the next two sections rules will be derived to avoid this problem. To reach that aim a graphical representation of the projected tensors due to Boerner will be employed. For two-column tableaux this representation, defined below, is easily seen to be in one-to-one correspondence with the representations by extended Rumer diagrams [39, 45], by bracket structures [38, 40] and by path diagrams [46, p. 30]. The equivalence of these representations has been discussed by Van Berkel [47].

Definitions.

1. Associate an index tableau $I^{[\lambda]}$ with a (possibly non-standard) Young tableau $T^{[\lambda]}$ and an index set $I = \{i_1, i_2, \ldots, i_N\}$, by replacing k in $T^{[\lambda]}$ by i_k, for k = 1, 2, ..., N.
2. An index tableau $I^{[\lambda]}$ obeying the following two rules is standard:
   (i) In the columns of $I^{[\lambda]}$ the indices increase from top to bottom.
   (ii) In the rows of $I^{[\lambda]}$ the indices do not decrease from left to right.
3. An index tableau that is standard, except for one or more columns containing equal indices, is half-standard.
4. An index tableau that is neither standard nor half-standard is called non-standard.

Examples.

1. $I = \{1, 7, 7, 8\}$

   $T^{[2,2]} = \begin{array}{|c|c|}\hline 3&2\\\hline 1&4\\\hline\end{array}$ $\Rightarrow I^{[2,2]} = \begin{array}{|c|c|}\hline 7&7\\\hline 1&8\\\hline\end{array}$ (a non-standard index tableau).

2. $I = \{1, 1, 2, 3\}$

   $T^{[2,1^2]} = \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array}$ $\Rightarrow I^{[2,1^2]} = \begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array}$ (a standard index tableau).

3. $I = \{4, 4, 5, 6\}$

   $T^{[3,1]} = \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2\\\cline{1-1}\end{array}$ $\Rightarrow I^{[3,1]} = \begin{array}{|c|c|c|}\hline 4&5&6\\\hline 4\\\cline{1-1}\end{array}$ (a half-standard index tableau).

Definition.

Let the index tableau $I^{[\lambda]}$ belong to the pair $\{I, T^{[\lambda]}\}$, then the tableau $\pi(I^{[\lambda]})$, $\pi \in S_N$, belongs to the pair $\{\pi(I), T^{[\lambda]}\}$.

Notes.

1. The action of π on $I^{[\lambda]}$ is given in the conventional way (sec. I.4.4) after numbering the positions of $I^{[\lambda]}$ according to the numbers in $T^{[\lambda]}$.
2. Recall that the action of a permutation on a Young tableau is defined by its action on the numbers in the tableau, not on the positions. This implies that the pair $\{I, T^{[\lambda]}\}$ yields the same index tableau as the pair $\{\pi(I), \pi(T^{[\lambda]})\}$.
3. The pair $\{I, \pi^{-1}(T^{[\lambda]})\}$ yields $\pi(I^{[\lambda]})$.

Example.

$I = \{i, j, k, \ell\}$, $T^{[\lambda]} = \begin{array}{|c|c|}\hline 3 & 4 \\ \hline 2 & 1 \\ \hline \end{array}$, $\pi = (123)$.

So:

$\pi(I) = \{k, i, j, \ell\}$, $\pi^{-1}(T^{[\lambda]}) = \begin{array}{|c|c|}\hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array}$, $\pi(T^{[\lambda]}) = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 3 & 2 \\ \hline \end{array}$.

Now:

$\{I, T^{[\lambda]}\}$ and $\{\pi(I), \pi(T^{[\lambda]})\}$ give $\begin{array}{|c|c|}\hline k & \ell \\ \hline j & i \\ \hline \end{array}$

and:

$\{\pi(I), T^{[\lambda]}\}$ and $\{I, \pi^{-1}(T^{[\lambda]})\}$ give $\begin{array}{|c|c|}\hline j & \ell \\ \hline i & k \\ \hline \end{array}$.

In the foregoing sections several operators associated with two Young tableaux were introduced. Henceforth these operators will be called tableau operators. The following tableau operators are of special importance:

(i) The partial antisymmetrizer $N^{[\lambda]}_{pk} \equiv \sigma^{[\lambda]}_{pk} N^{[\lambda]}_{k}$.

(ii) The Young unit $Y^{[\lambda]}_{pk}$.

(iii) The Y-Y unit $W^{[\lambda]}_{pk}$.

By means of an index tableau a tensor, resulting from $E_I \equiv v_{i_1} \otimes \ldots \otimes v_{i_N}$ by the action of a tableau operator, can be represented graphically.

Definition.

Let $I^{[\lambda]}_{\kappa}$ be the index tableau belonging to $\{I, T^{[\lambda]}_{k}\}$, then we define the following projected tensors.

(i) A partially antisymmetric tensor:

$$D^{(p)}_{I^{[\lambda]}_{k}} \equiv N^{[\lambda]}_{pk} E_I .$$

(ii) A Weyl-Rumer tensor:

$$F^{(p)}_{I^{[\lambda]}_{k}} \equiv Y^{[\lambda]}_{pk} E_I .$$

(iii) A Gelfand tensor:

$$G^{(p)}_{I^{[\lambda]}_{k}} \equiv W^{[\lambda]}_{pk} E_I .$$

Example.

Let the index tableau $\begin{array}{|c|c|}\hline j & i \\ \hline k & \ell \\ \hline \end{array}$ belong to $I = \{i, j, k, \ell\}$ and $T^{[2,2]}_{2} = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & 4 \\ \hline \end{array}$.

Then:

$$D^{(p)}_{\begin{array}{|c|c|}\hline j & i \\ \hline k & \ell \\ \hline \end{array}} \equiv N^{[2,2]}_{p2}\, E_{ijk\ell} = \sigma^{[2,2]}_{p2}\, N^{[2,2]}_{2}\, E_{ijk\ell}$$

$$= \sigma^{[2,2]}_{p2}\,[(1) - (23)][(1) - (14)]\, E_{ijk\ell}$$

$$= \sigma^{[2,2]}_{p2}\,(E_{ijk\ell} - E_{ikj\ell} - E_{\ell jki} - E_{\ell kji})$$

for certain p, $1 \leq p \leq 24$, numbering the Young tableaux, including the non-standard ones.

Notes.

1. The action of tableau operators onto tensor components can also be represented by index tableaux. Thus, for instance:

$$s^{I_k^{[\lambda]}}_{(p)} = Y^{[\lambda]}_{kp}\, t^{I}.$$

Here we let the first index of the tableau operator label the index tableau.

2. Note that in general:

$$\pi F^{(p)}_{I_k^{[\lambda]}} \neq F^{(p)}_{\pi(I_k^{[\lambda]})} \quad \text{and also} \quad \pi F^{(p)}_{I_k^{[\lambda]}} \neq F^{(p)}_{\pi^{-1}(I_k^{[\lambda]})},$$

but that:

$$\pi s^{I_k^{[\lambda]}}_{(p)} = s^{\pi^{-1}(I_k^{[\lambda]})}_{(p)},$$

which can be proved as follows (suppressing $[\lambda]$):

$$\pi s^{I_k}_{(p)} \equiv \pi Y_{kp}\, t^{I} = \sum_{P \in S_N} U(P)_{kp}\, \pi P\, t^{I} = \sum_{P} U(P)_{kp}\, P\, t^{\pi^{-1}(I)} = Y_{kp}\, t^{\pi^{-1}(I)}$$

$$\equiv s^{\pi^{-1}(I_k)}_{(p)}.$$

5 5 Weyl-Rumer basis

In this section it will be shown that the collection of Weyl-Rumer tensors labelled by standard index tableaux forms a complete and linearly independent basis of $V_n^{\otimes N}$. It seems that no earlier general proof of this result has been given, although the equivalent case of PN-projected tensors has been treated by Boerner [8] and independently also by Seligman [48]. The case

of NP-projected tensors is more difficult, however. Since the well-known proof of Rumer et al. [45] depends essentially on the fact that only one- and two-column tableaux are considered, this manner of proof cannot be generalized to general Young tableaux. The proof given in this section is an adaptation and extension of Boerner's proof.

Lemma.

(i) The partially antisymmetric tensor:

$$D^{(p)}_{I_k^{[\lambda]}}$$

is zero if two or more indices in one column of $I_k^{[\lambda]}$ are equal.

(ii) Every non-zero partially antisymmetric tensor is equal (up to sign) to a partially antisymmetric tensor with a column-ordered index tableau (this is an index tableau with the indices increasing in the columns from top to bottom).

(iii) The set of partially antisymmetric tensors with p, $[\lambda]$ and I fixed and k running over column-ordered index tableaux is linearly independent and contains non-zero tensors only.

Proof. Suppress superscript $[\lambda]$.

$$D^{(p)}_{I_k} \equiv N_{pk} E_I = N_p \sigma_{pk} E_I = N_p E_J = D^{(p)}_{J_p}$$

where we have written $J = \sigma_{pk}(I)$.

Note furthermore that $\{I, T_k\}$ and $\{\sigma_{pk}(I), \sigma_{pk}(T_k)\}$ both give I_k, and that $\{\sigma_{pk}(I), \sigma_{pk}(T_k)\} = \{J, T_p\}$, which gives J_p. So $I_k = J_p$. Because N_p is a product of antisymmetrizers, one for each column, the tensor $D^{(p)}_{J_p}$ is a product of antisymmetric tensors. Now the following points immediately follow from the standard theory of antisymmetric tensors (e.g. ref. 12, ch. VIII):

a. If, and only if, two or more equal indices occur in the same column of $J_p = I_k$ the tensor $D^{(p)}_{I_k}$ is zero.

b. Two partially antisymmetric tensors that are connected by a vertical permutation $\nu \in C_p$ are equal (up to the parity ζ_ν).

c. Two tensors which are not connected by a vertical permutation are linearly independent. (To make this plausible one can assume the basis of V_n to be orthogonal. The two tensors are then also orthogonal, and hence linearly independent).

The remainder of the proof follows trivially from these points.

Corollary 1.

The tensor space $V_n{}^{\otimes N}$ does not carry irreps of CS_N with more than n rows.

Proof. Express a Weyl-Rumer tensor in terms of partially antisymmetric tensors:

$$F^{(p)}_{I^{[\lambda]}_k} = Y^{[\lambda]}_{pk} E_I = N^{[\lambda]}_{pk} P^{[\lambda]}_k E_I = \sum_{\pi \in R^{[\lambda]}_k} N^{[\lambda]}_{pk} E_{\pi(I)} = \sum_{\pi \in R^{[\lambda]}_k} D^{(p)}_{\pi(I^{[\lambda]}_k)} .$$

If the number of different indices in I is less than the number of rows of $[\lambda]$, all index tableaux $\pi(I^{[\lambda]}_k)$, $\pi \in R^{[\lambda]}_k$, have necessarily two or more equal indices in one column; in that case the Weyl-Rumer tensor at the left is zero. The maximal number of different indices in any index set I labelling an element E_I of $V_n{}^{\otimes N}$ is n. This proves the corollary.

Note.

The irrep $[1^N]$ is the irrep of CS_N with the most rows in its Young diagram. Therefore, if N > n not all irreps of CS_N can be carried by $V_n{}^{\otimes N}$ and the representation $CS_N \rightarrow AS_N$ is non-faithful. (See page 150 of ref. 8 for a proof that the representation is faithful if $N \leq n$).

Corollary 2.

Let $[\lambda]$ have m columns, and let the index tableau $I^{[\lambda]}_k$ belong to $\{I, T^{[\lambda]}_k\}$. Then $F^{(p)}_{I^{[\lambda]}_k}$ is zero whenever any index in I occurs more than m times.

Proof. Express $F^{(p)}_{I^{[\lambda]}_k}$ again in partially antisymmetric tensors. Since one of the indices occurs more than m times, this index appears at least twice in one of the columns of all the partially antisymmetric tensors in the expansion. So they are all zero, and accordingly the W-R tensor vanishes.

Note.

These corollaries give rise to the Pauli principle in its original form, stating that an orbital may not be occupied by more than two fermions of intrinsic spin ½. We come back to this point in the next chapter.

Let U be a subspace of $V_n{}^{\otimes N}$ belonging to a certain index set I; that is, U is spanned by:

$$\{PE_I \mid \text{all } P \in S_N\}.$$

The space U, being stable under S_N, can be decomposed in the usual way into a direct sum of irreducible S_N carrier spaces. If $Y^{[\lambda]}_{qq} E_J \neq 0$, for some q and J, with $E_J \in U$, then:

$$\{Y^{[\lambda]}_{pq} E_J \mid p = 1, \ldots, f_{[\lambda]}\} \tag{1}$$

spans an irrep of S_N. Note that all elements of this basis are characterized by the same index tableau $J^{[\lambda]}_q$. So, this index tableau can serve as a label for the space spanned by this basis, and we write accordingly $U_{J^{[\lambda]}_q}$ for this subspace of U. By virtue of the following lemma this association of an irreducible carrier space of S_N with an index tableau is unique.

Lemma.

Let $I^{[\lambda]}_p$ belong to the pair $\{I, T^{[\lambda]}_p\}$ and $I^{[\lambda]}_q$ to $\{I, T^{[\lambda]}_q\}$, $p \neq q$. The Young tableaux $T^{[\lambda]}_p$ and $T^{[\lambda]}_q$ are not necessarily standard. Then:

$$I^{[\lambda]}_p = I^{[\lambda]}_q \text{ implies } F^{(k)}_{I^{[\lambda]}_p} = F^{(k)}_{I^{[\lambda]}_q},$$

for all k, $1 \leq k \leq N!$

Proof. Drop $[\lambda]$.

$\{I, T_q\}$ and $\{I, T_p\}$ give the same index tableau $I_q = I_p$. Also, since $\sigma^{-1}_{qp} = \sigma_{pq}$, $\{I, \sigma_{qp}(T_p)\}$gives the same index tableau as $\{\sigma_{pq}(I), T_p\}$, (see the preceding section). So $\{I, T_p\}$ ard $\{\sigma_{pq}(I), T_p\}$ both give I_p, hence $I = \sigma_{pq}(I)$.

Then.

$$F^{(k)}_{I_p} \equiv Y_{kp} E_I = Y_{kp} E_{\sigma_{pq}(I)} = Y_{kq} E_I = F^{(k)}_{I_q}.$$

Example.

$T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$, $T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$, $I = \{1, 2, 2\}$, $I_1 = I_2 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}$.

From the lemma:

$$Y_{11} E_{122} = Y_{12} E_{122} = F^{(1)}_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}}$$

$$Y_{21} E_{122} = Y_{22} E_{122} = F^{(2)}_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}}.$$

Notes.

1. This lemma lays the mathematical foundation under the use of index tableaux for labelling projected tensors. The same proof can be given for partially antisymmetric tensors; the same result will also be shown for Gelfand tensors in the next section.
2. The association of index tableaux and projected tensors with fixed superscript k is not one-to-one, as the converse of the lemma does not hold. For instance:

$$F^{(k)}_{I_p} = F^{(k)}_{\pi(I_p)}, \qquad \forall \pi \in R_p,$$

while in general: $I_p \neq \pi(I_p)$.

We are now ready to state and prove the main theorem of this section.

Theorem.

Let $U \subset V_n^{\otimes N}$ belong to the index set I. U can be decomposed into a direct sum of irreducible S_N carrier spaces:

$$U = \sum_{[\lambda]} \sum_{q} \oplus U^{[\lambda]}_{I_q},$$

where q runs over all possible standard index tableaux of shape $[\lambda]$ and index set I.

Proof. Without loss of generality U can be taken to be generated by an ordered index set I_o. That is, we assume U to be spanned by $\{PE_{I_o}\}$, $I_o = \{i_1 \leq i_2 \leq \ldots \leq i_N\}$.

The set of all W-R tensors characterized by standard or half-standard index tableaux forms a (generally overcomplete) basis of U. This follows since for any $E_{P(I_o)} \in U$ we have:

$$E_{P(I_o)} = (1)PE_{(I_o)} = \Big(\sum_{\lambda} \sum_{p=1}^{f_{[\lambda]}} \sum_{s=1}^{f_{[\lambda]}} \frac{f_{[\lambda]}}{N!} Y^{[\lambda]}_{ps}\, \eta^{[\lambda]}_{sp}\Big) PE_{(I_o)} = \sum_{\lambda} \sum_{p,q} c^{[\lambda]}_{qp}\, F^{(p)}_{I^{[\lambda]}_q},$$

where we introduced the resolution of identity afforded by Young units (sec. I.5.1) and used the fact that Young units span $f_{[\lambda]}$-dimensional right ideals (sec. I.5.1) and where we further introduced:

$$c^{[\lambda]}_{qp} = \sum_{s=1}^{f_{[\lambda]}} \frac{f_{[\lambda]}}{N!} R^{D^{[\lambda]}(P)}_{qs}\, \eta^{[\lambda]}_{sp}.$$

The index tableau $I^{[\lambda]}_q$ arises from $\{I_o, T^{[\lambda]}_q\}$. Since I_o is ordered and $T^{[\lambda]}_q$

is standard, this index tableau is either standard or half-standard.

Let us now assume that the index tableaux are ordered such that q first runs over the $g_{[\lambda]}$ standard index tableaux and then over the $f_{[\lambda]} - g_{[\lambda]}$ half-standard index tableaux of shape $[\lambda]$ and index set I_o. If we can show that the set:

$$\{F^{(p)}_{I_q^{[\lambda]}} \mid q = 1, \ldots, g_{[\lambda]};\ \text{fixed } p\} \tag{2}$$

forms a complete and linearly independent basis for the space $Y^{[\lambda]}_{pp}(U)$, we have proved the theorem. (Note that if we had used the PN-type Young units, completeness would already have been proved at this point, because in that case all half-standard tensors vanish).

The manner of proof is the following:

(i) First it is shown that a projected tensor $Y^{[\lambda]}_{pp} T$, arbitrary $T \in U$, does not have more than $g_{[\lambda]}$ linearly independent components, from which follows that the dimension of $Y^{[\lambda]}_{pp}(U)$ is less than or equal to $g_{[\lambda]}$.

(ii) Then it is shown that the set (2), which is manifestly a subset of $Y^{[\lambda]}_{pp}(U)$, is of dimension $g_{[\lambda]}$.

An arbitrary tensor $T \in U$ can be written as:

$$T = \sum_{Q \in S_N} t^{Q(I_o)} E_{Q(I_o)}$$

and the corresponding projected tensor (suppressing superscript $[\lambda]$) as:

$$Y_{pp}(T) = \sum_Q (Y_{pp}\, t^{Q(I_o)})\, E_{Q(I_o)} = \sum_Q s^{Q(I_p)}_{(p)}\, E_{Q(I_o)},$$

where $s^{Q(I_p)}_{(p)}$, the component of $Y_{pp}(T)$ along $E_{Q(I_o)}$, is labelled by the index tableau $Q(I_p)$. This index tableau can be standard, half-standard or non-standard. We now show that every projected tensor component can be expressed as a linear combination of the $g_{[\lambda]}$ components with standard index tableaux only. Suppressing subscript p, we write:

$$s^{Q(I_p)} = Y_{pp}\, t^{Q(I_o)} = Q^{-1} Y_{pp}\, t^{I_o} = \sum_{q=1}^{f_{[\lambda]}} L^{D(Q^{-1})}_{qp}\, Y_{qp}\, t^{I_o}$$

$$= \sum_{q=1}^{f_{[\lambda]}} L^{D(Q^{-1})}_{qp}\, s^{I_q}.$$

The index tableau I_q belongs to the ordered index set I_o and the standard

Young tableau T_q, hence I_q is either standard or half-standard. We now show that $s^{I_q} = 0$, if I_q is half-standard, that is, if I_q has two equal indices in one column. Let these indices be i_k and $i_{k'}$, and so $(k,k')I_q = I_q$.

Now:

$$s^{I_q} = s^{(k,k')I_q} = (k,k')s^{I_q} = -s^{I_q},$$

since (k,k') belongs to C_q and has parity -1.

Knowing that the half-standard components vanish, it follows that $s^{Q(I_p)}$ can be expressed in terms of standard components only. This proves point (1).

Remark: It may seem surprising that half-standard components vanish, whereas half-standard tensors do not. Notice that we have seen this before in an example (sec. I.4.4), where $Y_{11}E_{121} \equiv F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1\\\cline{1-1}\end{array}}$ and $Y_{11}t^{121} \equiv s^{\begin{array}{|c|c|}\hline 1&2\\\hline 1\\\cline{1-1}\end{array}}$ were explicitly worked out.

We now show that the $g_{[\lambda]}$ different standard Weyl-Rumer tensors F_{I_q} are linearly independent. We do this by expanding them in the linearly independent set of partially antisymmetric tensors D_{I_k}, introduced in the beginning of this section. First we order the standard W-R tensors in the following manner: Read down the columns of two standard index tableaux I_k and I_ℓ simultaneously, starting at the leftmost column. Let $i \in I_k$ and $j \in I_\ell$ be the first indices encountered that are different. Now if $i < j$, then F_{I_k} comes before F_{I_ℓ}, and vice versa.

Secondly we order the set $\{D_{I_k}\}$, such that the $g_{[\lambda]}$ standard tensors among them come first and are ordered in the same way as the standard W-R tensors. The ordering of the non-standard partially antisymmetric tensors is irrelevant. For the case $I_o = \{1, 2, 3, 4\}$ we have for instance the following ordering:

$$F_{\begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1}3\\\cline{1-1}\end{array}} < F_{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}4\\\cline{1-1}\end{array}} < F_{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}}$$

and

$$D_{\begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1}3\\\cline{1-1}\end{array}} < D_{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}4\\\cline{1-1}\end{array}} < D_{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}} < D_{\begin{array}{|c|c|}\hline 2&1\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}}$$

Using this ordering we shall show below that a set of linear equations of

the following general appearance holds:

$$(F_{I_1}, \ldots, F_{I_{g_{[\lambda]}}}) = (D_{I_1}, D_{I_2}, \ldots\ldots\ldots\ldots) \begin{pmatrix} 1 & & & \\ & 1 & & \underline{A} \\ & & \ddots & \\ & \underline{0} & & \ddots \\ & & & 1 \\ - & - & - & - \\ & & \underline{B} & \end{pmatrix}$$

The $g_{[\lambda]} \times g_{[\lambda]}$ submatrix in the upper part of the matrix is upper triangular and its diagonal elements are 1. Clearly the total matrix has $g_{[\lambda]}$ linearly independent columns, representing the standard W-R tensors in the basis $\{D_{I_k}\}$. So, the $g_{[\lambda]}$ standard W-R tensors are linearly independent, which proves the theorem.

It remains to show that the linear equations have this specific form. Expand to that end the standard W-R tensors (suppressing superscript p):

$$F_{I_k} = Y_{pk} E_{I_o} = \sum_{\pi \in R_k} N_{pk} E_{\pi(I_o)} = \sum_{\pi \in R_k} D_{\pi(I_k)}.$$

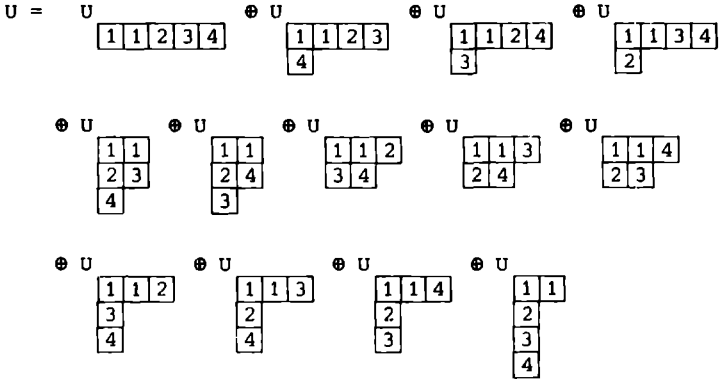
Replace the possibly occurring non-column-ordered tensors in the rightmost expression by column-ordered ones via:

$$D_{\pi(I_k)} = \zeta_\nu D_{\nu\pi(I_k)}, \qquad \nu \in C_k.$$

Look in the standard index tableau I_κ for the first column that changes under the permutation $\nu\pi$, and find in this column the highest element i that is replaced (by an element i'') under the action of $\nu\pi$. The following reasoning shows that i'' is necessarily larger than i The operation π replaces i by i', $i \leq i'$, ν replaces i' by an element i'' standing below i'. There are two possibilities: either i'' has not been moved by $\nu\pi$ and then immediately: $i'' > i$, or i'' has been placed there by π, and then it must have come from the right (remember we are looking at the leftmost column that changes) and so $i'' \geq i'$. Since $i' \geq i$, we have $i'' \geq i$. The tableau $\nu\pi(I_k)$ is either standard and then comes after I_k because $i'' > i$, or is non-standard and then comes after the last standard one. This proves the appearance of zeros below the diagonal of the upper part of the matrix. The diagonal elements finally follow from the fact that the identity (1) belongs to R_k, which implies that a standard W-R tensor F_{I_k} has the component +1 along the standard partially antisymmetric tensor D_{I_k} with the same index tableau I_k.

Example.

The index set {1, 1, 2, 3, 4} gives a 5!/2! dimensional space U, which decomposes under S_N as follows:



The space U has the dimension:

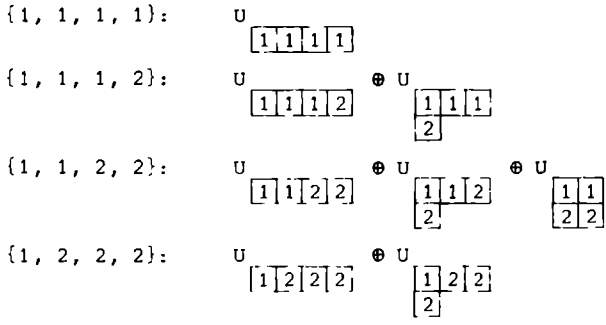
$$1\times1 + 3\times4 + 2\times5 + 3\times5 + 3\times6 + 1\times4 = 60.$$

The decomposition of the total tensor product space $V_n{}^{\otimes N}$ follows easily by letting the index set I run over all essentially different sets (that is, sets which do not follow from each other by permutation), and then decomposing the associated spaces according to the theorem just given. Clearly tensors belonging to different index sets are linearly independent and so we obtain a complete reduction of $V_n{}^{\otimes N}$ under S_N in this manner.

Example.

It is illustrative to look again at the decomposition of the four-electron spin space $V_2{}^{\otimes 4}$. In sec. I.4.7 we discussed this problem from the point of view of SU(2), we now approach from S_4.

The possible ordered index sets and the associated irreducible spaces are:



{2, 2, 2, 2}: U [2|2|2|2]

As in sec. I.4.7 we can assemble the basis vectors (standard W-R tensors) in the following table. From the theorem of sec. I.4.6 we know that these W-R tensors are eigenfunctions of the spin angular momentum operator S^2. Anticipating the association of the spin quantum number S with the shape $[\lambda]$ (chapter I.6), the spaces are labelled by S.

$V'_{S=2}$: $F_{[1|1|1|1]}$ $F_{[1|1|1|2]}$ $F_{[1|1|2|2]}$ $F_{[1|2|2|2]}$ $F_{[2|2|2|2]}$

$V'_{S=1}$: $F^{(1)}_{\begin{smallmatrix}1&1&1\\2\end{smallmatrix}}$ $F^{(1)}_{\begin{smallmatrix}1&1&2\\2\end{smallmatrix}}$ $F^{(1)}_{\begin{smallmatrix}1&2&2\\2\end{smallmatrix}}$

$V''_{S=1}$: $F^{(2)}_{\begin{smallmatrix}1&1&1\\2\end{smallmatrix}}$ $F^{(2)}_{\begin{smallmatrix}1&1&2\\2\end{smallmatrix}}$ $F^{(2)}_{\begin{smallmatrix}1&2&2\\2\end{smallmatrix}}$

$V'''_{S=1}$: $F^{(3)}_{\begin{smallmatrix}1&1&1\\2\end{smallmatrix}}$ $F^{(3)}_{\begin{smallmatrix}1&1&2\\2\end{smallmatrix}}$ $F^{(3)}_{\begin{smallmatrix}1&2&2\\2\end{smallmatrix}}$

$V'_{S=0}$: $F^{(1)}_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}}$

$V''_{S=0}$: $F^{(2)}_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}}$

For instance $F^{(1)}_{\begin{smallmatrix}1&1&2\\2\end{smallmatrix}}$ reads explicitly:

$$F^{(1)}_{\begin{smallmatrix}1&1&2\\2\end{smallmatrix}} \equiv Y_{11}^{[3,1]} E_{1122} = 2(E_{1122} - E_{2121}) + 2(E_{1212} - E_{2211})$$

$$= 2(\alpha\alpha\beta\beta - \beta\alpha\beta\alpha) + 2(\alpha\beta\alpha\beta - \beta\beta\alpha\alpha),$$

where the last line is in terms of the usual spin functions.

Notes.

1. In ref. 49, reprinted in this thesis, the graphical representation of W-R tensors is employed to classify the possible Russell-Saunders states of an atom.
2. Of course, now that we have decomposed $V_n{}^{\otimes N}$ under S_N, we have done the

same under GL(n), U(n), SU(n) and their respective Lie algebras. This follows from the theorem in sec. I.4.6.

5 6 Gelfand basis

Because Young units have a very simple appearance, a Weyl-Rumer basis of $V_n{}^{\otimes N}$ is easily constructed. However, the basis has the definite drawback of being non-orthogonal. On the other hand a Gelfand basis, obtained by projection with Y-Y units, is orthogonal, but has the disadvantage of being rather difficult to generate; for we have seen that a Y-Y unit contains in general N! permutations, all with a coefficient that cannot so easily be computed. Still, Gelfand bases find widespread application, not only in theoretical physics, but also in quantum chemistry.

The purpose of this section is to prove that the set of Gelfand tensors with standard index tableaux forms a complete linearly independent basis of $V_n{}^{\otimes N}$, just as is the case with W-R tensors.

Theorem 1.
The set Gelfand tensors:

$$\{G^{(p)}_{I_k^{[\lambda]}} \mid k \text{ running over standard tableaux, } p, [\lambda] \text{ fixed}\}$$

spans an irrep $<\lambda>$ of GL(n), U(n) and SU(n) and their respective Lie algebras. Letting furthermore $<\lambda>$ run over all diagrams with at most n rows and p from 1 to $f_{[\lambda]}$, the set thus obtained forms a complete and linearly independent basis of $V_n{}^{\otimes N}$.

Proof. Suppress superscripts on the Gelfand tensors.
Following an idea of Matsen [25] we employ the recursion relation for Y-Y units given in sec. I.5.2. By induction it will first be shown that the Gelfand tensor·

$$G_{I_p^{[\lambda]}} = W^{[\lambda]}_{pp}\, E_{I_o}, \qquad I_o = \{i_1 \leq i_2 \leq \ldots \leq i_N\}$$

is non-zero if, and only if, the index tableau $I_p^{[\lambda]}$ associated with $\{I_o, T_p^{[\lambda]}\}$ is standard. Further it will be proved that if the standard index tableau $I_p^{[\lambda]}$ associated with $\{I_o, T_p^{[\lambda]}\}$ is equal to the index tableau $I_k^{[\lambda]}$ associated with $\{I_o, T_k^{[\lambda]}\}$, $k \neq p$, then the corresponding Gelfand tensors are identical.

Remove, to prove these points, the parentheses from the inside out in the following expression:

$$W^{[\lambda]}_{pp}\, E_{I_o} = (W^{\lambda(N)} (W^{\lambda(N-1)} \ldots (W^{\lambda(1)} v_{i_1}) \otimes \ldots \otimes v_{i_{N-1}}) \otimes v_{i_N}).$$

Suppose that after removal of r pairs of brackets the following non-zero standard tensor has been obtained:

$$G_I \equiv G^{\lambda(r)}_{I_q} = (W^{\lambda(r)} (\ldots (W^{\lambda(1)} v_{i_1}) \otimes \ldots) \otimes v_{i_r})$$

$$= W^{\lambda(r)}_{qq} (v_{i_1} \otimes \ldots \otimes v_{i_r}).$$

Suppose also that the next s indices are equal:

$$i_r < i_{r+1} = i_{r+2} = \ldots = i_{r+s} < i_{r+s+1}$$

for certain s, $1 \leq s \leq N-r$. (Because in this induction process all equal indices are handled simultaneously, the possibility $i_{r+1} = i_r$ is excluded). Removing the next s pairs of brackets we in fact consider the construction of:

$$G^{\lambda(r+s)}_{I_{q'}} = W^{\lambda(r+s)}_{q'q'} (G_1 \otimes G_2)$$

where G_2 is the symmetric tensor $v_{i_{r+1}} \otimes \ldots \otimes v_{i_{r+s}}$. The tensor product $G_1 \otimes G_2$ belongs to an irrep $[\lambda(r)] \otimes [s]$ of $S_r \otimes S_s$. From it and its partners one can generate a space $\bar{V}$ by the action of the left coset generators of $S_r \otimes S_s$ in S_{r+s}. $\bar{V}$ carries the induced representation:

$$\Lambda = ([\lambda(r)] \otimes [s]) \uparrow S_{r+s},$$

which can be decomposed into a direct sum of irreps of S_{r+s} (ref. 4, p. 135). From the results in the appendix of ref. 44 it follows that $G_1 \otimes G_2$ has a component in every one of the irreducible S_{r+s}-carrier spaces contained in $\bar{V}$. So the tensor:

$$G^{\lambda(r+s)}_{I_{q'}}$$

is non-zero only if $[\lambda(r+s)]$ is contained <u>at least</u> once in Λ. The tensor is furthermore uniquely defined if $[\lambda(r+s)]$ occurs <u>just</u> once in Λ, which implies that if another Y-Y unit $W^{\lambda(r+s)}_{p'p'}$, $p' \neq q'$, projects also a non-zero tensor out of $G_1 \otimes G_2$, it gives the same result.

Invoking now Littlewood's rule [ref. 50, p. 92] we find that Λ contains all the irreps corresponding to Young diagrams that can be built by the addition of s boxes to $[\lambda(r)]$, no two boxes being added to the same column. All these irreps occur with multiplicity 1. So, if $[\lambda(r+s)]$

originates from $[\lambda(r)]$ in this manner, then the Gelfand tensor with an index tableau $I_{q'}^{\lambda(r+s)}$, obtained from $I_q^{\lambda(r)}$ by adding s boxes containing the equal indices $i_{r+1}, \ldots, i_{r+s}$ to its boundary, is non-vanishing and uniquely determined. We clarify this by an example:

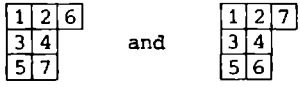
Let W be a Y-Y unit belonging to the standard Young tableau:

then the tensor

$$W(G_{\begin{array}{|c|c|}\hline i & j \\ \hline k & \ell \\ \hline \end{array}} \otimes G_{\begin{array}{|c|c|c|}\hline a & a & a \\ \hline \end{array}}) = G_{\begin{array}{|c|c|c|}\hline i & j & a \\ \hline k & \ell \\ \hline a & a \\ \hline \end{array}}$$

arises from E_{I_o}, where $I_o = \{i \leq j \leq k \leq \ell < a = a = a\}$.

Note that the Y-Y units belonging to the standard tableaux:



give identical results when acting on E_{I_o}.

Continuing the removal of brackets, one ends up with either a zero-tensor or a tensor with a standard index tableau. Varying p from 1 to $f_{[\lambda]}$ and $[\lambda]$ over all possible shapes one obtains a set standard Gelfand tensors, belonging to I_o. The tensors with different index tableaux are linearly independent by the construction just described; the tensors with the same index tableaux are identical. The partners in irreps of S_N of the tensors that are obtained in this way by the action of the diagonal Y-Y units can be generated by the corresponding off-diagonal Y-Y units. Retaining now only the S_N-carrier spaces characterized by different standard index tableaux we have obtained a complete and linearly independent basis for the space spanned by

$$\{E_{P(I_o)} \mid \text{all } P \in S_N\}.$$

Completeness follows immediately from the fact that the Y-Y units span right ideals. Varying finally I_o over all ordered index sets, a complete linearly independent basis of $V_n^{\otimes N}$ adapted to a decomposition under S_N is obtained.

Collecting all standard Gelfand tensors transforming as the k-th row of $[\lambda]$, $k = 1, \ldots, f_{[\lambda]}$, we obtain by the result of sec. I.4.6 a basis for the irrep $\langle\lambda\rangle$ shared by GL(n), U(n) and SU(n). This basis carries also irreps of the respective Lie algebras (sec. I.3.1).

Often Gelfand tensors are represented [30, 31] in a manner which visualizes their behaviour upon subduction under the chain.

$$GL(n) \supset GL(n-1) \supset \ldots \supset GL(1).$$

In order to be consistent, we derive this well-known representation from the definition given earlier (sec. I.5.4), and to that end we need Weyl's branching theorem [ref. 6, p. 391]. It is convenient to denote from here on the irreps of GL(n) by partition symbols $\langle\mu\rangle$ in which all the n row lenghts are given explicitly, even the ones of length zero.

Theorem 2 (Weyl's branching theorem).

Upon subduction to GL(n-1) the irrep $\langle\mu_n\rangle$ of GL(n) decomposes in the following manner:

$$\langle\mu_n\rangle \equiv \langle m_{1n}, m_{2n}, \ldots, m_{nn}\rangle = \sum_{t=0}^{N} \sum_{\langle\mu_{n-1}\rangle} \oplus (\langle\mu_{n-1}\rangle \otimes \langle t\rangle),$$

where:

$\langle t\rangle$ is the symmetric and only tensor representation of GL(1).

$\langle\mu_{n-1}\rangle = \langle m_{1,n-1}, m_{2,n-1}, \ldots, m_{n-1,n-1}\rangle$ runs over all partitions of N-t, subject to the constraint: $m_{j+1,n} \geq m_{j,n-1} \geq m_{jn}$ for $j = 1, 2, \ldots, n-1$. The partition $\langle\mu_{n-1}\rangle$ denotes an irrep of GL(n-1).

Proof. Formula 29 of ref. 51 yields for this case:

$$\langle\mu_n\rangle = \sum_{t=0}^{N} \sum_{\langle\mu_{n-1}\rangle} \oplus g_{\mu_n\mu_{n-1}t} (\langle\mu_{n-1}\rangle \otimes \langle t\rangle),$$

where we have written $g_{\mu_n\mu_{n-1}t}$ for the multiplicity of $\langle\mu_{n-1}\rangle \otimes \langle t\rangle$ in $\langle\mu_n\rangle$, and $\langle\mu_{n-1}\rangle$ runs over all partitions of N-t. This follows because $\langle t\rangle$ is the only tensor representation of GL(1). The multiplicity factors are the same as the ones occurring in the induction·

$$([\mu_{n-1}] \otimes [t]) \uparrow S_N = \sum_{[\mu_n]} \oplus g_{\mu_n\mu_{n-1}t} [\mu_n]$$

as also has been discussed in ref. 51.

Invoking Littlewood's rule [50, p. 92] it follows that.

$g_{\mu_n\mu_{n-1}t} = 1$ if $[\mu_n]$ can be constructed from $[\mu_{n-1}]$ by addition of t boxes to $[\mu_{n-1}]$ without adding more than one box to a column.

$g_{\mu_n\mu_{n-1}t} = 0$ in all other cases.

It is now easy to see that $g_{\mu_n\mu_{n-1}t} = 1$ if:

$$m_{j+1,n} \geq m_{j,n-1} \geq m_{jn}$$

for all $j = 1, \ldots, n-1$.

Example.

Subduction to GL(3) of the irrep $\langle 3,2,1,0\rangle$ of GL(4) carried by $V_4^{\otimes 6}$:

$$\langle 3,2,1,0\rangle = (\langle 3,2,1\rangle) \otimes \langle 0\rangle \oplus (\langle 2,2,1\rangle \oplus \langle 3,1,1\rangle \oplus \langle 3,2,0\rangle) \otimes \langle 1\rangle$$
$$\oplus (\langle 2,1,1\rangle \oplus \langle 2,2,0\rangle \oplus \langle 3,1,0\rangle) \otimes \langle 2\rangle \oplus (\langle 2,1,0\rangle) \otimes \langle 3\rangle.$$

In ref. 44 it is exhibited how to construct an element in a basis for $\langle \mu_n \rangle$ that is simultaneously an element in a basis for $\langle \mu_{n-1} \rangle \otimes \langle t \rangle$. The procedure is the following: first adapt a tensor $E_{i_1 i_2 \ldots i_{N-t}}$ to $[\mu_{n-1}]$ of S_{N-t}, where $1 \leq i_j \leq N-1$ and $j = 1, \ldots, N-t$. This yields G_I, associated with $I_{q'}^{[\mu_{n-1}]}$. Then adapt:

$$G_I \otimes \underbrace{v_n \otimes v_n \otimes \ldots \otimes v_n}_{t \text{ factors}}$$

to S_N, yielding a Gelfand tensor G, characterized by $I_q^{[\mu_n]}$. But this is exactly the manner in which the Gelfand tensors were constructed during the proof of theorem 1, and hence Gelfand tensors are sequence-adapted to $GL(n) \supset GL(n-1)$, that is, they carry simultaneously $\langle \mu_n \rangle$ and $\langle \mu_{n-1} \rangle \otimes \langle t \rangle$. The partition $\langle \mu_{n-1} \rangle$ is obtained by stripping all t boxes containing the index n off $I_q^{[\mu_n]}$. Repeating this process for n, n-1, n-2, ..., 1, noting that all accompanying subductions are multiplicity free and that the tail group GL(1) has only one-dimensional tensor irreps, that is, the chain $GL(n) \supset GL(n-1) \supset \ldots \supset GL(1)$ is canonical, we find that Gelfand tensors can be uniquely labelled by the successive partitions: $\langle \mu_n \rangle, \langle \mu_{n-1} \rangle, \ldots, \langle \mu_1 \rangle$. It is common to write these partitions in the form of a Gelfand pattern:

$$G \equiv W_{qq}^{[\mu_n]} E_I \equiv \begin{pmatrix} m_{1n} & m_{2n} & m_{3n} & \cdots\cdots\cdots & m_{nn} \\ m_{1,n-1} & m_{2,n-1} & \cdots\cdots\cdots & m_{n-1,n-1} & \\ m_{1,n-2} & \cdots\cdots & & & \\ & & & & \\ & & m_{11} & & \end{pmatrix}$$

Recall again that Weyl's branching theorem states: $m_{i+1,j+1} \geq m_{ij} \geq m_{i,j+1}$.

Example.

Element in V_6 $\otimes^9$:

G $\equiv$

1	2	2
3	4	
4	5	
5		
6		

$$\begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 0 \\ & 3 & 2 & 2 & 1 & 0 \\ & & 3 & 2 & 1 & 0 \\ & & & 3 & 1 & 0 \\ & & & & 3 & 0 \\ & & & & & 1 \end{pmatrix}$$

Notes.

1. Clearly a Gelfand tensor is labelled uniquely by the genealogy $\langle \mu_n \rangle$, $\langle \mu_{n-1} \rangle$, ..., $\langle \mu_1 \rangle$. This follows directly from the chain GL(n) $\supset$ GL(n-1) $\supset$ ... $\supset$ GL(1) being canonical. As pointed out in sec. I.4.7 an irrep $\langle \mu_k \rangle$ can alternatively be labelled by the eigenvalues of k Casimir invariants. If we do this for k = n, n-1, ..., 1, then every Gelfand tensor can also be labelled uniquely by the eigenvalues of the $\frac{1}{2}n(n+1)$ invariants associated with this canonical chain, or in other words: these Casimir invariants form a complete set of commuting operators.
2. The fact that projection with Y-Y units, which themselves are adapted to the canonical chain $S_N \supset S_{N-1} \supset \ldots \supset S_1$, yields tensors adapted to the chain GL(n) $\supset$ GL(n-1) $\supset$ ... $\supset$ GL(1) is of course well-known. For example it encompasses also the result that genealogically constructed spin functions are identical with Y-Y projected spin functions [46, app. 2].

5 7 Explicit construction of matrix representations of GL(n)

In the computer calculations, reported on in the second part of this dissertation, actual irreducible tensor representations of GL(n) have been employed to study the effect of a basis transformation in the one-electron space V_n onto the Weyl-Rumer tensors. In this section it will be discussed how the irreducible matrices representing the elements of GL(n) may be constructed. Naturally these matrices depend on the choice of basis for $V_n \otimes^N$.

Let us first consider the basis projected by the Y-Y units $W_{kk}^{[\lambda]}$, because in this case the computation is straightforward. The following notation will be used throughout this section:

$E'_I = (\alpha v_{i_1}) \otimes (\alpha v_{i_2}) \otimes \ldots \otimes (\alpha v_{i_N})$, $\alpha \in GL(n)$.

(a^i_j) is the matrix of α.

$A^I_J = a^{i_1}_{j_1} a^{i_2}_{j_2} \ldots a^{i_N}_{j_N}$.

I_o, J_o are ordered index sets (indices not decreasing from left to right).

$C(I) = 1/(N_1!\, N_2! \ldots N_N!)$,

where N_k is the number of times i_k occurs in I, k = 1, ..., N.

Furthermore the superscript $[\lambda]$ will be omitted everywhere in this section.

The rows and the columns of the matrix $\underline{T}(\alpha)$, representing α on an irreducible subspace of $V_n^{\otimes N}$, are labelled by standard index tableaux, just as the Gelfand basis itself. If I_k is a standard index tableau belonging to the index set I_o and the standard Young tableau T_k (see sec. I.5.4), then one can write:

$$G'_{I_k} = W_{kk}(E'_{I_o}) = W_{kk}\Big(\sum_J A^J_{I_o} E_J\Big) = W_{kk}\Big(\sum_{J_o} C(J_o) \sum_{P \in S_N} A^{P(J_o)}_{I_o} E_{P(J_o)}\Big),$$

where the summation over all possible index sets J has been broken up into one over ordered index sets J_o and a sum over all the permutations of J_o. The factor $C(J_o)$ corrects for the presence of equal indices in J_o. Using:

$$W_{kk} E_{P(J_o)} = W_{kk} P E_{J_o} = \sum_q O(P)_{kq} W_{kq} E_{J_o} = \sum_q O(P)_{kq} G_{J_q},$$

where $\underline{O}(P)$ is the orthogonal Young-Yamanouchi matrix (sec. I.5.3), (recall that the Y-Y units span minimal right ideals), and:

$$A^{P(J_o)}_{I_o} = P^{-1} A^{J_o}_{I_o},$$

(sec. I.4.4), one gets:

$$G'_{I_k} = \sum_{J_o} C(J_o) \Big(\sum_q \Big(\sum_{P \in S_N} O(P)_{kq} P^{-1} \Big) A^{J_o}_{I_o} G_{J_q} \Big).$$

Further defining:

$$A^{J_q}_{I_k} \equiv \sum_{P \in S_N} O(P)_{qk} A^{P^{-1}(J_o)}_{I_o} = \frac{N!}{f_{[\lambda]}} W_{qk} A^{J_o}_{I_o}$$

one gets:

$$G'_{I_k} = \sum_{J_o} C(J_o) \sum_q A^{J_q}_{I_k} G_{J_q}.$$

The summation over q in this expression is still running over all $f_{[\lambda]}$

standard Young tableaux, and so J_q is either a standard or a half-standard index tableau. However, it has been shown in the preceding section that the half-standard Gelfand tensors are zero and that the ones with the same standard index tableau are identical. So, we may let q run over the different standard index tableaux only, if we introduce a factor $m_q(J_o)$, which is the number of times J_q occurs in the original sum over q. Finally one arrives at:

$$G'_{I_k} = \sum_{J_o} \sum_{q} C(J_o)\, m_q(J_o)\, A^{J_q}_{I_k}\, G_{J_q} .$$

The prescription for the calculation of $\underline{T}(\alpha)$ is now transparent: The columns of $\underline{T}(\alpha)$ are labelled by the different standard index tableaux I_k, which are obtained by letting I_o run over all ordered index sets and k over all the standard index tableaux associated with I_o. The rows, characterized by J_q, are similarly obtained from letting J_o run over all ordered index sets and q over all different standard index tableaux associated with J_o. The matrix element

$$T(\alpha)^{J_q}_{I_k} \equiv C(J_o)\, m_q(J_o)\, A^{J_q}_{I_k}$$

follows by projection with $\frac{N!}{f_{[\lambda]}} W_{qk}$ onto $A^{J_o}_{I_o}$.

Note.

This prescription is equivalent to the one given by Littlewood [50, p. 183 ff] for the construction of "invariant matrices". Littlewood employs to this end natural units of the PN-type.

The tensor representations carried by the W-R basis are more difficult to derive. When trying to follow the procedure above, one finds that the half-standard tensors do not drop out of the expressions, because they are in general non-zero, as we have seen in sec. I.5.5. So, we must first be able to express the half-standard tensors in terms of the standard ones. Although it is not possible to give explicit relations, one can derive sets of linear equations from which the half-standard tensors can be obtained in terms of the standard tensors.

Let (p,q) be the permutation that interchanges the two equal indices that are by definition present in one of the columns of the half-standard index tableau I_j. Then we get the equation:

$$F_{I_J} = F_{(p,q)I_J} = Y_{JJ}\ (p,q)\ E_I = \sum_k R^{D(p,q)}{}_{kJ}\ Y_{Jk}\ E_I = \sum_k R^{D(p,q)}{}_{kJ}\ F_{I_k},$$

where $\underline{R}^{D(p,q)}$ is the matrix representation of (p,q), introduced in sec. I.5.1. As we have seen there, this matrix has a very simple appearance; and consequently the half-standard tensors can be easily eliminated from equations of this type.

Example.

Consider $[\lambda] = [2^2, 1]$ and $I = \{1, 1, 2, 3, 4\}$.

The linear equations generated by $(p,q) = (12)$ read:

$$(F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&3\\\hline 4\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&4\\\hline 3\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&3\\\hline 1&4\\\hline 2\\\cline{1-1}\end{array}}) = (F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&3\\\hline 4\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&4\\\hline 3\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&3\\\hline 4\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&4\\\hline 3\\\cline{1-1}\end{array}}\ F_{\begin{array}{|c|c|}\hline 1&3\\\hline 1&4\\\hline 2\\\cline{1-1}\end{array}})\begin{pmatrix} -1 & 0 & 1\\ 0 & -1 & 1\\ -1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1 \end{pmatrix}$$

These equations yield the following expressions for the half-standard tensors:

$$F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&3\\\hline 4\\\cline{1-1}\end{array}} = -\tfrac{1}{2}F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&3\\\hline 4\\\cline{1-1}\end{array}},\quad F_{\begin{array}{|c|c|}\hline 1&2\\\hline 1&4\\\hline 3\\\cline{1-1}\end{array}} = -\tfrac{1}{2}F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&4\\\hline 3\\\cline{1-1}\end{array}},\quad F_{\begin{array}{|c|c|}\hline 1&3\\\hline 1&4\\\hline 2\\\cline{1-1}\end{array}} = \tfrac{1}{2}F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&3\\\hline 4\\\cline{1-1}\end{array}} + \tfrac{1}{2}F_{\begin{array}{|c|c|}\hline 1&1\\\hline 2&4\\\hline 3\\\cline{1-1}\end{array}}$$

We are dealing with sets $\{F_{I_k} \mid k = 1, \ldots, f_{[\lambda]}\}$ in which only the tensors labelled by different standard index tableaux are linearly independent. Tensors characterized by the same (standard) index tableau are identical, half-standard tensors in $\{F_{I_k}\}$ depend linearly on the standard tensors. We now choose a maximal linearly independent subset of $\{F_{I_k}\}$, and label its elements by r, s, t, ... The remaining tensors in $\{F_{I_k}\}$ are labelled by a, b, c, ...; the whole set is labelled by i, j, k, ... We can then summarize the foregoing considerations in the one formula:

$$F_{I_a} = \sum_r S(I)_{ra}\ F_{I_r}.$$

The matrix $\underline{S}(I)$ can be computed in the manner just described and is in the sequel assumed to be known.

The computation of the tensor representation carried by the W-R basis is now straightforward:

$$F'_{I_r} = \sum_{J_o} C(J_o) \sum_{\imath} \{ \sum_{P \in S_N} R^{D(P)}{}_{\imath r} A_{I_o}^{P(J_o)} \} F_{J_\imath}.$$

Recall (sec. I.5.1):

$$R^{D(P)} = \underline{X}^{-1} \underline{U}(P^{-1})$$

and so the expression between square brackets becomes:

$$\sum_{P} \sum_{\jmath} \eta_{\imath\jmath} U(P^{-1})_{\jmath r} A_{I_o}^{P(J_o)} = \sum_{\jmath} \eta_{\imath\jmath} Y_{\jmath r} A_{I_o}^{J_o},$$

where $\eta_{\imath\jmath}$ is the $(\imath,\jmath)$-matrix element of $\underline{X}^{-1}$. Now, $Y_{\jmath r} A_{I_o}^{J_o} = 0$ if $J_\jmath$ is half-standard (see the proof of the main theorem of sec. I.5.5) and the sum over $\jmath$ can accordingly be replaced by one over s, if we correct for the $m_s(J_o)$-fold occurrence of the standard index tableau J_s.
Defining:

$$A_{I_r}^{J_\imath} = \sum_{s} m_s(J_o) \eta_{\imath s} Y_{sr} A_{I_o}^{J_o}$$

we get:

$$F'_{I_r} = \sum_{J_o} C(J_o) \sum_{\imath} A_{I_r}^{J_\imath} F_{J_\imath}.$$

The summation over $\imath$ in this expression runs still over all standard and half-standard index tableaux. J_o runs over all ordered index sets that have no indices occurring more often than either the number of rows or the number of columns of $[\lambda]$, because if any index appears with a frequency higher than either of these two numbers $F_{J_\imath}$ is zero.

Inserting the expressions derived earlier, one finally arrives at

$$F'_{I_r} = \sum_{J_o} C(J_o) \sum_{t} (A_{I_r}^{J_t} + \sum_{a} S(J_o)_{ta} A_{I_r}^{J_a}) F_{J_t},$$

from which one concludes that the (J_t, I_r)-matrix element is

$$T(\alpha)_{I_r}^{J_t} = C(J_o) [\sum_{s} m_s(J_o) (\eta_{ts} + \sum_{a} S(J_o)_{ta} \eta_{as}) Y_{sr}] A_{I_o}^{J_o}.$$

This expression looks rather complicated, but usually simplifies drastically in practice, because most of the times η_{as} is equal to zero for all values of its row and column index, which run over the half-standard and standard index tableaux respectively. Indeed, I have not been able to find a one- or two-columned representation with an η_{as}-value not equal to zero for any index set. Note further that η_{ts} is almost always equal to δ_{ts}

and we see that the calculation of $T(\alpha)^{J_t}_{I_r}$ usually amounts to not much more than projection with Y_{tr}.

To clarify these remarks we turn to the example $[2^2, 1]$, treated in sec. I.5.1. The matrix $\underline{X}^{-1}$ is here:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The first row is labelled by:

$$T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 \\ \cline{1-1} \end{array}$$

The only off-diagonal element not equal to zero is η_{15}. The ordered index sets $J_o = \{j_1 \leq j_2 \leq j_3 \leq j_4 \leq j_5\}$ that give half-standard index tableaux in cooperation with T_1 satisfy:

$$j_1 = j_3 \quad \text{and/or} \quad j_2 = j_4 \quad \text{and/or} \quad j_3 = j_5 \quad \text{and/or} \quad j_1 = j_5 .$$

But then, because of the ordering:

$$j_1 = j_2 = j_3 \text{ and/or } j_2 = j_3 = j_4 \text{ and/or } j_3 = j_4 = j_5 \text{ and/or } j_1 = j_2 = j_3 = j_4 = j_5 .$$

And so J_o has at least one index triply occurring, and hence does not contribute to tensor representations with two-columned diagrams. Therefore in this example η_{15} does not contribute to $\underline{T}(\alpha)$ and $\underline{S}(J)$ does not arise for any index set J.

Note.

Very recently another formula for the tensor representation carried by the W-R basis has been given by Brunet and Seligman [52]. Their derivation is completely different from the one just given. Also their final formula has an entirely different appearance.

5 8 A note on the tensor irreps of the Lie algebra of GL(n)

In this section a few of the Lie algebraic connotations of the decomposition of tensor space will be discussed. First we introduce some general definitions concerning the representations of Lie algebras. Since our main

interest is in GL(n), these definitions are all given for the Lie algebra AL(n) of GL(n). However, these definitions also hold for arbitrary Lie algebras; see ref. 16 for further details.

Definitions.

1. Let E^{ij}, i, j = 1, ..., n be the linear operator on V_n with the matrix:

$$(\underline{E}^{ij})_{k\ell} \equiv \delta^i_k \, \delta^j_\ell .$$

As discussed in sec. I.3.2, these n^2 operators span AL(n), the Lie algebra of GL(n).

2. Let $\mathcal{H}_n$ be the n-dimensional subalgebra of AL(n) spanned by:

$$\{E^{ii} \mid i = 1, \ldots, n\}.$$

It is easy to show that this algebra is the maximal commutative subalgebra of AL(n): the Cartan subalgebra of AL(n).

3. Let D be an m-dimensional linear Lie algebra representation of AL(n). Let D be carried by V_m. Since the elements of the Cartan subalgebra $\mathcal{H}_n$ commute, all the linear operators D(H), $H \in \mathcal{H}_n$, can be simultaneously diagonalized on V_m. Let $v_o \in V_m$ be a common eigenvector of the elements of $D(\mathcal{H}_n)$, that is.

$$D(H)\ v_o = \alpha(H)\ v_o, \qquad \forall\ H \in \mathcal{H}_n,\ 0 \neq v_o \in V_m,\ \alpha(H) \in C.$$

The set of complex eigenvalues $\alpha(H)$ can be considered as a mapping of $\mathcal{H}_n$ into C. This mapping α is linear, for:

$$D(\lambda H + \lambda' H')\ v_o = \lambda D(H)\ v_o + \lambda' D(H')\ v_o = \lambda\alpha(H)\ v_o + \lambda'\alpha(H')\ v_o,$$

and also:

$$D(\lambda H + \lambda' H')\ v_o \equiv \alpha(\lambda H + \lambda' H')\ v_o .$$

Hence:

$$\alpha(\lambda H + \lambda' H') = \lambda\alpha(H) + \lambda'\alpha(H')$$

with $\lambda, \lambda' \in C$, $H, H' \in \mathcal{H}_n$.

The linear functional $\alpha: \mathcal{H}_n \to C$, associated with the common eigenvector $v_o \in V_m$, is called a weight of the representation D.

4. Let $H = \sum_i h_i E^{ii}$ belong to $\mathcal{H}_n$, then:

$$\alpha(H) = \sum_i h_i\, \alpha(E^{ii}) = \sum_i h_i\, \alpha_i$$

with $\alpha_i \equiv \alpha(E^{ii})$.

So, every eigenvalue $\alpha(H)$ can be expressed in terms of $\alpha_1, \alpha_2, \ldots, \alpha_n$. These scalars constitute a weight vector, associated with the

representation D and the eigenvector $v_o \in V_m$. (The word vector is often omitted in this definition).

5. As we will see below, the components of the weight vectors associated with tensor representations of AL(n) are natural numbers. It is convenient to order these weight vectors in the following manner:
The weight vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ comes before the weight vector $(\alpha_1', \alpha_2', \ldots, \alpha_n')$, if the first non-zero difference $\alpha_i - \alpha_i'$ $(i = 1,\ldots,n)$ is positive. The weight $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is then called <u>higher</u> than the weight $(\alpha_1', \alpha_2', \ldots, \alpha_n')$; so it makes sense to speak of the <u>highest weight</u> of a tensor representation.

We proceed by showing that an arbitrary Weyl-Rumer or Gelfand tensor is a simultaneous eigenvector of all elements in the Cartan subalgebra of AL(n).

In sec. I.4.3 the following tensor representation T of AL(n) carried by $V_n^{\otimes N}$ has been defined:

$$T: E^{ij} \rightarrow \sum_{k=1}^{N} E^{ij}(k).$$

The operator E^{ij} acts on a basis vector v_p of V_n in the following manner:

$$E^{ij} v_p = \delta_p^j v_i,$$

where v_i is also a basis vector of V_n. Now:

$$T(E^{jj})\ E_I \equiv \sum_{k=1}^{N} v_{i_1} \otimes \ldots \otimes v_{i_{k-1}} \otimes (E^{jj} v_{i_k}) \otimes \ldots \otimes v_{i_N} = N_j E_I,$$

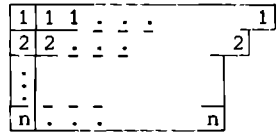
where the "occupation number" N_j is the number of times the index j occurs in the index set I which labels the tensor E_I. Physically, the operator $T(E^{jj})$ represents the numbers of particles in orbital v_j and is therefore called a <u>number operator</u>. The element E^{jj} is by definition a basis element of $\mathcal{H}_n \subset AL(n)$, and so the set of occupation numbers $(N_1, N_2, \ldots, N_n)$ is a weight vector associated with the tensor representation T and the tensor $E_I \in V_n^{\otimes N}$. Every tensor E_I is a common eigenvector of the n number operators $T(E^{jj})$, $j = 1, \ldots, n$. Clearly, all elements of the subspace U of $V_n^{\otimes N}$ spanned by:

$$\{E_{P(I)} \mid \forall P \in S_N, \text{ fixed } I\}$$

also belong to this same weight (or in more physical terms: belong to the same configuration I). Adaptation of U to S_N by means of either Young units or Y-Y units does not change the weight of the elements of U, and

hence we conclude that all W-R and Gelfand tensors, arising from E_I, belong to the weight vector $(N_1, N_2, \ldots, N_n)$.

Consider the tensor irrep $\langle \lambda_1, \lambda_2, \ldots, \lambda_n \rangle$ of AL(n). A W-R or Gelfand tensor with the following standard index tableau:



where the i-th row has λ_i boxes (i = 1, ..., n), is associated with weight $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. It can easily be verified that no standard index tableau of this shape can be constructed that is associated with a weight higher than $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ (see sec. I.5.5 or I.5.6). So, the row lengths of the Young diagram are the components of the highest weight vector associated with the irrep $\langle \lambda_1, \lambda_2, \ldots, \lambda_n \rangle$ of AL(n). Since a tensor irrep of AL(n) is uniquely characterized by a Young diagram it is also determined unambiguously by the components of the highest weight vector.

The elements $T(E^{ij})$, $i < j$, change a tensor of weight:

$$(N_1, N_2, \ldots, N_i, \ldots, N_j, \ldots, N_n)$$

into one of weight:

$$(N_1, N_2, \ldots, N_i+1, \ldots, N_j-1, \ldots, N_n).$$

Because the new weight is higher, $T(E^{ij})$ is often called a <u>raising operator</u>. Similarly $T(E^{ij})$, $i > j$, is a <u>lowering operator</u>.

One can prove that $E'_I \equiv T(E^{ij})E_I$ $(i \neq j)$ is of weight N_i+1 either by inspection or by using: $T(E^{ij}) = T(E^{ii})T(E^{ij}) - T(E^{ij})T(E^{ii})$ from which follows $T(E^{ii})E'_I = (N_i+1)E'_I$. Similarly one can prove that E'_I is of weight N_j-1 by using: $T(E^{ij}) = T(E^{ij})T(E^{jj}) - T(E^{jj})T(E^{ij})$.

<u>Example.</u>

The Gelfand tensor $G_{\begin{smallmatrix}1&1&1&2\\2&2\end{smallmatrix}}$ is an element in a basis for $\langle 4,2 \rangle$ of AL(2), $\Lambda(U(2))$ and $\Lambda(SU(2))$. It has weight (3,3). The raising operator $T(E^{12})$ has the following effect:

$$T(E^{12})G_{\begin{smallmatrix}1&1&1&2\\2&2\end{smallmatrix}} = G_{\begin{smallmatrix}1&1&1&1\\2&2\end{smallmatrix}} + G_{\begin{smallmatrix}1&1&1\\1&2\end{smallmatrix}\,1\,2} + G_{\begin{smallmatrix}1&1&1&2\\2&1\end{smallmatrix}} = G_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}\,1\,1\,1}.$$

The resulting tensor is of weight (4,2).

The Casimir invariants of U(2) and SU(2) can easily be expressed in terms of the raising and lowering operators. For instance in the case of SU(2) one first goes over to a traceless basis of $\Lambda(SU(2))$, via:

$$\overline{E}^{12} \equiv E^{12} \equiv \sigma_+/2$$

$$\overline{E}^{21} \equiv E^{21} \equiv \sigma_-/2$$

$$\overline{E}^{11} \equiv (E^{11} - E^{22})/2 \equiv \sigma_z/2$$

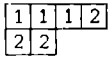
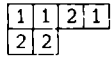
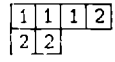
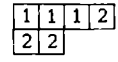
$$\overline{E}^{22} \equiv -(E^{11} - E^{22})/2 = -\sigma_z/2.$$

Note that this basis differs by a simple linear transformation from the one introduced in the example of sec. I.3.4.
The Casimir invariants are then [31]:

$$I_1^{(2)} = \overline{E}^{11} + \overline{E}^{22} = 0$$

$$I_2^{(2)} = \overline{E}^{11}\overline{E}^{11} + \overline{E}^{12}\overline{E}^{21} + \overline{E}^{21}\overline{E}^{12} + \overline{E}^{22}\overline{E}^{22} = 2S^2.$$

From the theory of chapter I.4 we know that Gelfand tensors are eigenvectors of the Casimir operators (the representations of the Casimir invariants). Using these expressions for the invariants one can easily calculate the eigenvalues:

$T(I_2^{(2)})G$  $= 2(G$  $+ G$  $) = 2(1(1+1))G$ 

The reader is referred to the work of Biedenharn et al. [53], Louck [31], Moshinsky [54] and others for the representation theory of SU(n) in this manner, which is in fact a generalization to SU(n) of the Racah-Wigner calculus of SU(2).

CHAPTER I 6 PHYSICAL APPLICATIONS

In this chapter the mathematics introduced in the preceding chapters will be applied to the physics of N-particle systems. In particular it will be shown that GL(n) is a dynamical group and S_N a symmetry group of a model system satisfying the following conditions:

- The system contains N identical particles, all moving in a stationary outer electric field (e.g. electrons moving in the field of one or more clamped nuclei).
- Magnetic and relativistic (including spin) contributions to the energy are absent. (These contributions can be introduced in a later stage as perturbations).
- The Hilbert space of the system is $[(2\sigma+1) \times n]^N$ dimensional, where σ is the spin of the particles and n is finite.

The great majority of the quantum chemical calculations performed to date satisfy these conditions.

In order to separate the approximating assumptions about the form of the wave function from the customary exact quantum mechanical postulates, the first two sections of this chapter will be devoted to a brief review of the axioms regarding the form of exact N-particle wave functions.

. 6 1 N-particle state vectors

The chief purpose of this section is to establish some terminology and to introduce the notation used in later sections.

Definitions.

1. A Euclidean space V is a vector space furnished with a positive definite inner product (see e.g. ref. 55). Henceforth this inner product will be denoted by $\langle f|g \rangle$, $f, g \in V$. A positive definite inner product induces a metric [55] on V, and so every Euclidean space is a metric space. Euclidean spaces are also called pre-Hilbert, unitary or inner product spaces.
2. A metric space is complete if every Cauchy sequence [20, p. 45] converges to an element of the space.
3. A Hilbert space is a complete Euclidean space.
4. A function $f: R^m \to C$ is square integrable if the integral:

$$\int_{\text{all } R^m} f^*(x) f(x)\,dx$$

is well-defined and finite.

5. The class of all square integrable functions on R^m will be denoted by $L^2(R^m)$.

Examples.

1. A Euclidean space of finite dimension is complete and hence a Hilbert space. (Completeness follows almost immediately from the completeness of the real line [56, p. 55]).
2. The parameter space of U(n) is a complete (non-linear) subspace of R^{n^2} with respect to the usual metric of R^{n^2}.
3. Endowing $L^2(R^m)$ with the inner product:

$$\langle f|g\rangle \equiv \int_{\text{all } R^m} f^*(x) g(x)\,dx$$

the set becomes a Hilbert space (see theorem II.4.3 and II.4.4 of ref. 55).

Notes.

1. Defining the norm of $f \in L^2(R^m)$ by

$$|f| = \langle f|f\rangle^{\frac{1}{2}}$$

it follows from the definition that all elements of $L^2(R^m)$ have a finite norm.

2. All Hilbert spaces of interest to us have a countable orthonormal basis that is complete (in the vector space sense of the term). Hilbert spaces having such a basis are called separable.

Definitions.

1. Let V be a Euclidean space. One can define an inner product on $V \otimes^N$ thus:

$$\langle f_1 \otimes \ldots \otimes f_N \mid g_1 \otimes \ldots \otimes g_N\rangle = \langle f_1|g_1\rangle \ldots \langle f_N|g_N\rangle.$$

It is easily seen that this inner product is positive definite, and so $V \otimes^N$ is a Euclidean space.

2. Let V be a Hilbert space. If V is of infinite dimension, $V \otimes^N$ is not necessarily complete. However, the completion of $V \otimes^N$ is uniquely determined [55, p. 144] and will also be denoted by $V \otimes^N$. From here on the definition of the tensor product is assumed to be extended in such a

manner that $V \otimes^N$ is always a Hilbert space.

Note.

Interpreting a tensor product of functions as a product function in the ordinary sense, the Hilbert space $L^2(R^3) \otimes^N$ is imbedded in $L^2(R^{3N})$. It can be proved that:

$$L^2(R^3) \otimes^N = L^2(R^{3N}).$$

(See ref. 55, p. 144).

Definition.

Let v be an arbitrary fixed vector in the linear space V. The set

$$\{av \mid \text{all } a \in C\}$$

is a ray in V.

After having introduced this necessary mathematical terminology we are now ready to introduce a physics postulate.

Postulate.

The state of a system consisting of N identical particles of spin σ is uniquely characterized by a ray in the Hilbert space:

$$\mathcal{H} \equiv L^2(R^{3N}) \otimes (L^2(s) \otimes^N).$$

Here:

$L^2(R^{3N})$ is the spatial Hilbert space of square integrable functions on R^{3N}.

$L^2(s)$ is the spin Hilbert space of dimension $2\sigma+1$.

Notes.

1. Instead of considering rays it is customary to take a representative of the ray by normalizing the state to unity. Even so a phase factor of absolute magnitude 1 is left unspecified.
2. The system Hilbert space can be written as the tensor product of one particle Hilbert spaces:

$$\mathcal{H} = [L^2(R^3) \otimes L^2(s)] \otimes^N.$$

3. A vector in $L^2(R^{3N})$ will be written as $\Phi(\vec{r})$ with $\vec{r} = (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \in R^{3N}$. Analogously a vector in $L^2(s) \otimes^N$ written as $\Theta(s_1, s_2, \ldots, s_N)$, where s_i is the spin coordinate particle i, i = 1, ..., N.

4. Note that $L^2(R^3) \otimes L^2(s)$ is isomorphic with the space of $(2\sigma+1)$-component spinors (ref. 4, p. 192). This isomorphism also preserves inner products. So, the postulate above is equivalent to one which expresses that a state is characterized by a ray in a tensor product of the Hilbert spaces of $(2\sigma+1)$-component spinors.
5. Although every state of a system is characterized by a ray in a Hilbert space, the converse is not true; that is, not every ray in a Hilbert space, associated with a system, represents a realizable state of this system.

6 2 Symmetry of N-particle systems

Before turning our attention to systems consisting of N identical particles, we first review briefly the case of one particle.

Definitions.

1. The set of all unitary and antiunitary operators [4, p. 203] on an n-dimensional space is denoted by $U^{\pm}(n)$. It is easy to show that $U^{\pm}(n)$ forms a group.
2. Let $V_n = V_{n_1} \oplus V_{n_2}$. The set of all unitary and antiunitary operators on W that leave both V_{n_1} and V_{n_2} invariant is denoted by $U^{\pm}(n_1) \times U^{\pm}(n_2)$. This set forms a group.

Consider a one-particle system with configuration space R^3 and Hamiltonian $h(\vec{r})$, $\vec{r} \in R^3$. Assume that $L^2(R^3)$ is completely decomposed into eigenspaces of $h(\vec{r})$:

$$h(\vec{r}) \mid \varepsilon_i, k_i \rangle = \varepsilon_i \mid \varepsilon_i, k_i \rangle, \qquad i = 1, 2, \ldots; \; k_i = 1, \ldots, n_i.$$

Here k_i runs over the degenerate states of energy ε_i. In this equation as in the rest of this chapter, it is ignored that part of the spectrum of $h(\vec{r})$ may be continuous. All operators on $L^2(R^3)$ that commute with $h(\vec{r})$ and that are either unitary or antiunitary form a group G_h^f, the full commutator group of $h(\vec{r})$. Knowing the degeneracies n_i, $i = 1, 2, \ldots$, of $h(\vec{r})$, it can easily be derived [57] that

$$G_h^f = U^{\pm}(n_1) \times U^{\pm}(n_2) \times U^{\pm}(n_3) \times \ldots$$

And so each eigenspace of $h(\vec{r})$ is invariant and irreducible under G_h^f. Clearly the full commutator group of $h(\vec{r})$ lacks physical meaning and is

not helpful in any way to the simplification of practical calculations. The group is only known after all degeneracies of $h(\vec{r})$ have been given and hence cannot contribute to an a priori knowledge of the splitting of energy levels.

However, if physical arguments, e.g. of geometric or Lie algebraic nature, make it possible to find subgroups of G_h^f, then group theory will become useful. If a subgroup of G_h^f can be found that shares with G_h^f all the irreps carried by $L^2(R^3)$, or in other words if a subgroup can be found that predicts the spectrum of $h(\vec{r})$, then group theory contributes to the solution of quantum mechanical problems in the maximal possible way. In general G_h^f can have several such spectrum predicting subgroups, the smallest of these will be called the symmetry group of $h(\vec{r})$, denoted by G_h^s.

Simple geometric arguments enable us almost always to define at least one physically meaningful subgroup of G_h^f: the invariance group G_h^1 of $h(\vec{r})$. This group is associated with the maximal group of isometric operators on R^3 that leave $h(\vec{r})$ invariant. (See e.g. ref. 4, p. 214 for the association of a Hilbert space operator with an R^3 operator). Since isometric operators on R^3 give rise to unitary Hilbert space operators and since the elements of G_h^1 commute with $h(\vec{r})$ [4, p. 214], it follows that the invariance group is a subgroup of G_h^f. Also, because the eigenspaces of $h(\vec{r})$ are stable under G_h^1, it follows that the invariance group is a (proper or improper) subgroup of the symmetry group.

Example.

To illustrate these points we consider the hydrogen atom. The energy levels are labelled by n = 1, 2, ..., the n-th level is n^2-fold degenerate, and so the full commutator group is $U^{\pm}((1)^2) \times U^{\pm}((2)^2) \times \ldots$. It has been found by Fock [58] that the 4-dimensional rotation-reflection group O(4) predicts the (discrete part of the) spectrum of the H-atom, and so O(4) is the symmetry group. Finally, the invariance group is O(3) (this group does of course not predict the degeneracy of levels with the same n- but different ℓ-quantum numbers). Summarizing:

$$U^{\pm}(1) \times U^{\pm}(4) \times U^{\pm}(9) \times \ldots \supset O(4) \supset O(3).$$

In this discussion of one-particle symmetry the definition of the symmetry group of the Hamiltonian $h(\vec{r})$ lacked preciseness; and indeed one can only be assured that G_h^s really predicts the spectrum when this is known

completely. In most of the few cases where the spectrum is known exactly one finds degeneracies that are higher than those that can be expected on the grounds of the invariance group alone. Besides the H-atom, other well-known cases showing "hidden" symmetry are the 3-dimensional harmonic oscillator and the spherical top [59]. Also many model Hamiltonians have this feature [60].

Nevertheless, the one particle symmetry group G_h^S is assumed to be known in the now following discussion of N-particle symmetry.

Definition.

Consider a system of N identical particles with Hamiltonian

$$H = \sum_{i=1}^{N} h(i) + \sum_{i<j}^{N} h(i,j)$$

acting on the Hilbert space $\mathcal{H}$. Let G_h^S be the symmetry group of $h(i)$, and let G_h be the maximal (proper or improper) subgroup of G_h^S which is such that the elements of the N-fold inner direct product group (sec. I.4.2) $[G_h]^N$ commute with H. The symmetry group G_H of H is defined by:

$$G_H \equiv [G_h]^N \otimes S_N,$$

where S_N permutes the spatial and spin coordinates of the particles, i.e. the particle labels.

Notes.

1. The elements of $[G_h]^N$ commute with those of S_N; this justifies the use of the direct product symbol in the definition of G_H.
2. Elements of G_H are unitary or antiunitary. This follows because the elements of S_N are unitary and the elements of G_h^S are either unitary or antiunitary.
3. Unitarity or antiunitarity of an element $U \in G_H$ assures that the transformed eigenstates $U|E\rangle$ of H cannot be distinguished by any measurement from the original states $|E\rangle$.
4. The symmetry group G_H contains operators which effect the same symmetry transformation on each particle separately, followed by a permutation of the identical particles among themselves. (Such a permutation is evidently an N-particle symmetry operator).
5. The occurrence of an inner direct product in the definition of G_H is necessary if we wish to interpret the elements of Lie algebras of

symmetry groups as observables. For if this specific form of G_H were not stipulated we would be led to consider Lie algebras of outer direct product (or perhaps even larger) groups. The elements of such algebras do not commute with S_N, so they either do not correspond to observables, or they can distinguish between the different particles in contradiction with the assumption that they are identical.

6. The symmetry group G_H may or may not be identical with $[G_h^S]^N \otimes S_N$. A spinless N-electron atom yields an example of a proper imbedding: the N-electron symmetry group is $S_N \otimes [O(3)]^N$, while the one-particle symmetry group is O(4). The appearance of O(3) rather than O(4) in the inner product is due to the presence of electron-electron interaction terms in H. In general, the interaction term h(i,j) may or may not lead to a symmetry group in which G_h is a proper subgroup of G_h^S.

The following postulate gives G_H quantum mechanical meaning.

Irreducibility Postulate.

The eigenstates of H belonging to the same eigenvalue of H carry an irrep of G_H.

Knowledge of G_H and its Hilbert space irreps enables the prediction of a good deal about the spectral decomposition of H. All degeneracies can be predicted and labelled. The transformation properties of the eigenstates of H can be derived, even if the states themselves are not known. From these one can derive e.g. selection rules, and rules to describe the splittings of levels in perturbing fields, etc. All this and much more has been worked out in detail in many text books on group theory. So, we forgo any discussion about these general aspects of the irreducibility postulate.

The symmetry group G_H is a direct product of the two commuting groups $[G_h]^N$ and S_N. To obtain a decomposition of the system Hilbert space

$$\mathcal{H} = [L^2(R^3) \otimes L^2(s)]^{\otimes N}$$

under G_H one can use this fact by proceeding in the following manner

(1) First decompose $\mathcal{H}$ into irreducible S_N-carrier spaces. All elements transforming according to the k-th row of the irrep $[\lambda]$ span a space $\mathcal{H}_k^{[\lambda]} \subset \mathcal{H}$. This space is invariant under $[G_h]^N$. (Apply, to prove this, Schur's lemma in the manner of the proof given in sec. I.2.3).

(ii) Decompose the spaces $\mathcal{H}_k^{[\lambda]}$, $k = 1, \ldots, f_{[\lambda]}$ under $[G_h]^N$, applying within each space the same basis transformation.

Let

$$\mathcal{H}_{k,p}^{(\lambda,\mu)} \subset \mathcal{H}_k^{[\lambda]}$$

carry the representation (μ) of $[G_h]^N$. The index p labels the multiple occurrence of (μ) in $\mathcal{H}_k^{[\lambda]}$. It now immediately follows that

$$V_p^{(\lambda,\mu)} \equiv \sum_{k=1}^{f_{[\lambda]}} \oplus \mathcal{H}_{k,p}^{(\lambda,\mu)}$$

carries an irrep of $G_H \equiv S_N \otimes [G_h]^N$.

If one would wish to arrive at a complete decomposition of $\mathcal{H}$ into eigenspaces of H, then this is as far as group theory can go. To obtain a complete spectral decomposition of H one can continue as follows: Take from every space $V_p^{(\lambda,\mu)}$, $p = 1, 2, \ldots$, the basis element that transforms to the same row of the irrep (λ,μ) of G_H, say the q-th row. The set:

$$\{\,|[\lambda], (\mu), q, p\rangle \mid p = 1, 2, \ldots\}$$

then spans a subspace of $\mathcal{H}$ invariant under H. Diagonalize H over this set. The irreducibility postulate states that no eigenvalue of H occurs more than once in this space, and so these eigenvalues may replace p to index uniquely the basis vectors. The index q labels the degenerate states.

The preceding discussion exhibits that the eigenstates of H are labelled by (μ) and $[\lambda]$. The following postulate, however, shows that $[\lambda]$ is superfluous for physically realizable states.

Pauli Postulate.

Only states belonging to one-dimensional irreps of S_N are physically realizable. Systems of identical particles with integral spin belong to the symmetric representation, systems of identical particles with half-integral spin belong to the antisymmetric representation.

Notes.

1. The Pauli principle has essentially the form of a superselection rule [61, p. 32], with for instance the Hermitean class sum operators of S_N (which are non-observable constants of motion), playing the rôle of superselection operators.
2. The link between spin and permutation symmetry, here introduced as a

postulate, has been shown by Pauli [62] to be unavoidable in a consistent quantum field theory.

3. The Pauli postulate can be replaced by another postulate, viz. the permutations of identical particles cannot be observed. See ref. 63 for a discussion of this alternative.

6 3. A model for N-electron systems

In this section some restrictions and approximations will be introduced leading to a widely used model for N-electron systems.

One of the assumptions defining this model is:

M1. There are no spin-dependent contributions to the energy. Consequently the Hamiltonian has the form:

$$H = H(\vec{r}) \otimes 1(\vec{s}),$$

where $H(\vec{r})$ is an energy operator on $L^2(R^3)^{\otimes N}$ and $1(\vec{s})$ is the unit operator on $L^2(s)^{\otimes N}$.

Notes.

1. From here on we will ignore that H may have antiunitary symmetry operators. This non-essential simplification is introduced as it would go beyond the scope of this work to explain where the presence of antiunitary symmetry operators causes doubling of degeneracies. See ref. 4 for a careful discussion of the cases where this occurs.
2. The assumption M1 and the simplification introduced in the foregoing note imply that the symmetry group G_H has the form:

$$G_H = [G_h^{(r)} \otimes SU(2)]^N \otimes S_N ,$$

where $G_h^{(r)} \otimes SU(2)$ is a subgroup of the one-particle space-spin symmetry group G_h^s, introduced in the foregoing section. SU(2) arises here because every unitary operator on the 2-dimensional space $L^2(s)$ commutes with the one-electron part of H. It has been pointed out in sec. I.4.6 that it makes no difference whether we consider U(2) or SU(2). According to common practice we have chosen the group with the smallest number of parameters.

The second assumption is in general a rather strong approximation, but is usually necessary for calculations on systems with more than one electron

and other systems where the Schrödinger equation cannot be solved exactly.

M2a. The Hilbert space $\mathcal{H}_M$ of the model system is finite dimensional:

$$\mathcal{H}_M = V_m \otimes (L^2(s) \otimes^N),$$

where V_m is an appropriately chosen m-dimensional subspace of $L^2(R^{3N})$.

This assumption can be formulated alternatively thus:

M2b. Replace everywhere the unit operator on $\mathcal{H}$ by the projector onto $\mathcal{H}_M$.

Let $|k\rangle$, $k = 1, 2, \ldots, \infty$ be an orthonormal basis of $\mathcal{H}$, and let the first M vectors span $\mathcal{H}_M$, $M = m \times 2^N$. The assumption M2b then reads:

$$1_\infty = \sum_{k=1}^{\infty} |k\rangle\langle k| \sim 1_M = \sum_{k=1}^{M} |k\rangle\langle k|.$$

We obtain a representation Q_M of an observable Q in the following manner:

$$Q \equiv 1_\infty Q 1_\infty \sim 1_M Q 1_M = \sum_{k,\ell=1}^{M} |k\rangle\langle k|Q|\ell\rangle\langle \ell| \equiv Q_M.$$

The homomorphism condition for the multiplication of observables is only approximately fulfilled:

$$(PQ)_M \equiv 1_M (PQ) 1_M = (1_M P 1_\infty)(1_\infty Q 1_M) \sim (1_M P 1_M)(1_M Q 1_M) = P_M Q_M.$$

So strictly speaking Q_M is not a representation of Q.

Further it is easy to show that $Q = Q^\dagger$ implies $Q_M = Q_M^\dagger$.

In practical calculations one often solves the eigenvalue problem of H_M in order to obtain approximations to eigenstates of the exact Hamiltonian H. The following theorem provides a theoretical foundation for this procedure.

Theorem.

Let the following eigenvalue equations hold:

$$H|E_k\rangle = E_k|E_k\rangle \quad \text{and} \quad H_M|E_k^{(M)}\rangle = E_k^{(M)}|E_k^{(M)}\rangle .$$

Then:

(i) $\lim_{M\to\infty} E_k^{(M)} = E_k$, and the sequence $\{E_k^{(M)} | M = 1, 2, \ldots\}$ is monotonically decreasing.

(ii) $\lim_{M\to\infty} \left| |E_k^{(M)}\rangle - |E_k\rangle \right| = 0.$

Proof. First replace the eigenvalue equations by variational equations and then invoke the variational principle. See ref. 64, p. 415 ff. for the details of the proof.

How good an approximation one gets by diagonalizing H_M is solely determined by the choice of the model Hilbert space $\mathcal{H}_M$. This choice has been the subject of many computational experiments. In these experiments one nearly always chooses $\mathcal{H}_M$ to be of the form.

$$\mathcal{H}_M = [V_n \otimes L^2(s)] \otimes^N,$$

where V_n is an n-dimensional subspace of the one-particle Hilbert space (orbital space) $L^2(R^3)$. Of course V_n must satisfy some basic requirements. For instance it must belong to the domain of the observables of interest. Thus, in order to be able to compute the kinetic energy of the electrons, the elements of V_n must be at least twice differentiable. If, for instance, the eigenstates of H_M have to reflect the symmetry of the exact solutions, V_n must be invariant under the one-particle symmetry group $G_h^{(r)}$, etc.

A model which uses a tensor product of finite-dimensional orbital spaces as the system Hilbert space is often called a multi-configuration model. Examples are provided by such well-known quantum chemical methods as the "limited" configuration interaction method or the "full" configuration interaction method. In both cases V_n is spanned by self-consistent field molecular orbitals; in the limited case only part of the basis of $\mathcal{H}_M$ is employed, whereas in full configuration interaction the complete basis of $\mathcal{H}_M$ is used. Also the valence bond method (V_n is spanned by atomic orbitals) and the multi-configuration self-consistent field method (orbitals and N-electron states are optimized simultaneously) belong to the category of multi-configuration methods.

From here on we will only consider model Hilbert spaces of the form just introduced:

$$\mathcal{H}_M = [V_n \otimes L^2(s)] \otimes^N.$$

Let H_M be a spin-free operator on this space, it is then easy to show that:

$$H_M = H_M(\vec{r}) \otimes 1(\vec{s}),$$

where $H_M(\vec{r})$ is an energy operator on $V_n \otimes^N$ and $1(\vec{s})$ is the unit operator on the spin space. $H_M(\vec{r})$ contains the coordinates of the N identical electrons constituting the system and commutes with all operators permuting these coordinates, so $H_M(\vec{r})$ is bisymmetric. In sec. I.4.5 it has been proved that the algebra of bisymmetric operators on $V_n \otimes^N$ is generated by $\Lambda([GL(n)]^N)$, therefore $H_M(\vec{r})$ is expressible as a polynomial in the generators of this algebra. The following formula, in this form due to Matsen

[25], gives this relation explicitly:

$$H_M(\vec{r}) = \sum_{i,j=1}^{n} H_{ij}^{(1)} [\sum_{k=1}^{N} E^{ij}(k)] +$$

$$+ \tfrac{1}{2} \sum_{i,j=1}^{n} \sum_{p,q=1}^{n} H_{ij;pq}^{(2)} [\sum_{k,\ell=1}^{N} E^{ip}(k)\, E^{jq}(\ell) - \delta_{pj} \sum_{k=1}^{N} E^{iq}(k)].$$

To explain the meaning of the symbols we sketch part of the derivation. Let V_n be spanned by the orthogonal basis $\{|i\rangle \mid i = 1, \ldots, n\}$, and let the spatial part of the exact Hamiltonian be:

$$H(\vec{r}) = \sum_{k=1}^{N} H^{(1)}(\vec{r}_k) + \tfrac{1}{2} \sum_{k,\ell=1}^{N} \{H^{(2)}(\vec{r}_k, \vec{r}_\ell) - \delta_{k\ell} H^{(2)}(\vec{r}_k, \vec{r}_\ell)\}.$$

Write the unit operator on $V_n^{\otimes N}$ as:

$$1_M = \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} |i_1 i_2 \cdots i_N\rangle\langle i_1 i_2 \cdots i_N|,$$

where $|i_1 i_2 \cdots i_N\rangle$ stands for $|i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_N\rangle$.
Then the one-particle part becomes:

$$H_M^{(1)}(\vec{r}) \equiv 1_M\, H^{(1)}(\vec{r})\, 1_M = \sum_{k=1}^{N} \{(\sum_{i_1=1}^{n} |i_1\rangle\langle i_1|) \otimes (\sum_{i_2=1}^{n} |i_2\rangle\langle i_2|) \otimes \cdots$$

$$\cdots \otimes (\sum_{i_k=1}^{n} \sum_{j_k=1}^{n} |i_k\rangle\langle i_k| H^{(1)} |j_k\rangle\langle j_k|) \otimes \cdots \otimes (\sum_{i_N=1}^{N} |i_N\rangle\langle i_N|)\} =$$

$$= \sum_{i,j=1}^{n} H_{ij}^{(1)} [\sum_{k=1}^{N} 1 \otimes 1 \otimes \cdots \otimes \underbrace{(|i\rangle\langle j|)}_{k\text{-th position}} \otimes \cdots \otimes 1] = \sum_{i,j=1}^{n} H_{ij}^{(1)} [\sum_{k=1}^{N} E^{ij}(k)]$$

We have used:

$$\langle i|j\rangle = \delta_{ij}$$

$$\sum_{i=1}^{n} |i\rangle\langle i| = 1$$

$$H_{ij}^{(1)} \equiv \langle i|H^{(1)}|j\rangle$$

and:

$$E^{ij}(k) \equiv 1 \otimes 1 \otimes \cdots \otimes |i\rangle\langle j| \otimes \cdots \otimes 1.$$

Clearly the operator $|i\rangle\langle j|$ has the matrix $\underline{E}^{ij}$, defined in sec. I.2.2 as a basis element of the algebra AL(n,C), which is the Lie algebra of GL(n), as we have seen in sec. I.3.2. In sec. I.4.3 it has been exhibited that $\Lambda([GL(n)]^N)$ is spanned by elements of the form:

$$\sum_{k=1}^{N} E^{ij}(k).$$

Hence the one-electron part of $H_M(\vec{r})$ is a linear function of the generators of the Lie algebra of $[GL(n)]^N$.

In exactly the same way one proceeds for the two-electron part of $H_M(\vec{r})$. Defining: $H^{(2)}_{ij;pq} \equiv \langle ij|\frac{1}{r_{12}}|pq\rangle$ one gets the final result. Note that the two-electron part is a second order polynomial in the generators of $\Lambda([GL(n)]^N)$. The fact that $H_M(\vec{r})$ is such a simple expression follows from the fact that the basis of V_n was taken to be orthogonal.

The model operator $H_M(\vec{r})$ being bisymmetric has also the interesting consequence that GL(n) is a dynamical group of the model system.

Definition.

A dynamical group of a system is a Lie group with a Lie algebra containing sufficient generators to express the Hamiltonian of the system as a function of the generators.

Well-known systems with a dynamical group are the three-dimensional harmonic oscillator [13, ch. 10] and the hydrogen atom [65]. In these two cases the dynamical group happens to be also the symmetry group of the system. This has the interesting consequence that the eigenvalues and eigenvectors of these systems can be determined from group theoretical arguments alone, without ever having to solve the Schrödinger equation explicitly.

To exhibit how the knowledge can help us that GL(n) is a dynamical group (not a symmetry group!) of our model system, we assume that $V_n \otimes^N$ has been decomposed into orthogonal irreducible GL(n)-carrier spaces. Assemble again all basis vectors carrying the irrep $\langle\lambda\rangle$ in one array.

$$\begin{array}{cccc}
\phi^{\langle\lambda\rangle}_{11} & \phi^{\langle\lambda\rangle}_{12} & \cdots\cdots\cdots\cdots & \phi^{\langle\lambda\rangle}_{1n_{\langle\lambda\rangle}} \\
\phi^{\langle\lambda\rangle}_{21} & \cdots\cdots\cdots & & \vdots \\
\vdots & & & \vdots \\
\vdots & & & \vdots \\
\phi^{\langle\lambda\rangle}_{f_{[\lambda]}1} & \cdots\cdots\cdots\cdots\cdots\cdots & & \phi^{\langle\lambda\rangle}_{f_{[\lambda]}n_{\langle\lambda\rangle}}
\end{array}$$

By definition the rows span identical irreps of GL(n). (From the result of sec. I.4.6 it follows that the columns span identical irreps of S_N). Each row spans also an irrep of the algebra of bisymmetric operators (sec. I.4.5)

and is hence invariant under $H_M(\vec{r})$. Since the different rows are orthogonal, the spaces are non-interacting under $H_M(\vec{r})$, and so this model operator can be diagonalized within each row separately. If we know the explicit representation of the generators of GL(n) carried by the subspaces, we can also write down the explicit representation of the model Hamiltonian on basis of each of the rows. This is the approach to the multi-configuration problem described in ref. 34 and 35.

Notes.

1. We are not yet in a position to consider which irreducible representations of GL(n) are physically realizable. This will be discussed in the next section, where we will find that only irreps with not more than $(2\sigma+1)$ columns are allowed for half-integral σ.
2. Although, as we have seen, the labelling of eigenstates of H according to irreps of S_N is superfluous since only antisymmetric representations of S_N are physically realizable, the labelling according to irreps of $S_N^{(r)}$, the permutation group acting on spatial coordinates alone, is non-trivial. This justifies the attention paid to the representation theory of S_N in the preceding chapter.
3. Much work [31, 53, 54] has been done to generalize the Wigner-Eckart theorem, so that the matrix elements of the raising and lowering operators E^{ij} can be expressed in terms of the vector coupling coefficients of GL(n). This seems a promising approach to the N-electron problem.

6 4 Symmetry restrictions imposed by spin

In this section the same physical model will be considered as in the preceding section, but in a somewhat different manner. Whereas in the second part of the foregoing section the attention was restricted to the spatial part $H_M(\vec{r})$ of the total spin-free Hamiltonian $H_M \equiv H_M(\vec{r}) \otimes 1(\vec{s})$, and accordingly only the spatial part of the model Hilbert space was considered, spin will remain included in the derivations presented subsequently. This will enable us to show that, although spin does not contribute to the dynamics of the model system, it still plays an important rôle by imposing certain symmetry restrictions. In particular we will be able to identify the physically realizable irreps of $S_N^{(r)}$.

For the following exposition the spatial one-particle symmetry is inessential; for reasons of clarity we assume it to be absent. The symmetry group of the system is then:

$$G_H = [1(\vec{r})]^N \otimes [SU(2)]^N \otimes S_N,$$

where $1(\vec{r})$ is the unit operator on $L^2(R^3)$. Since group theory states that inequivalent irreducible carrier spaces of the symmetry group are non-interacting under the Hamiltonian of the system, we do not introduce an extra approximation, when H_M is diagonalized with respect to a basis of an irreducible subspace of the model Hilbert space

$$\mathcal{H}_M = (V_n^{\otimes N}) \otimes (L^2(s)^{\otimes N}).$$

To construct such a basis we first adapt $L^2(s)^{\otimes N}$ to SU(2) and then adapt the whole spatial-spin Hilbert space to S_N; that is, we subsequently anti-symmetrize.

In the foregoing chapters it has been discussed how to adapt a tensor product space to SU(n). We recapitulate the main findings, specialized to spin space:

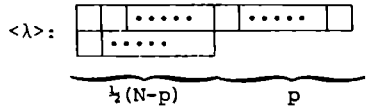
- $L^2(s)^{\otimes N}$ can be decomposed employing Lie algebraic techniques, e.g. ladder operators or vector coupling (sec. I.4.7). This decomposition under $\Lambda(SU(2))$ induces a complete decomposition under $S_N^{(s)}$, the permutation group acting on the spin coordinates (sec. I.4.6).
- Alternatively, $L^2(s)^{\otimes N}$ can be decomposed under $S_N^{(s)}$. This induces a decomposition into irreducible SU(2)-carrier spaces (sec. I.5.5 and I.5.6).
- Each irreducible SU(2)-carrier space $V^{<\lambda>}$ in $L^2(s)^{\otimes N}$ can be labelled by a partition $<\lambda>$ of N.
- Each space $V^{<\lambda>}$ occurs $f_{[\lambda]}$ times, where $f_{[\lambda]}$ is the dimension of the corresponding irrep $[\lambda]$ of $S_N^{(s)}$.
- The direct sum $V^{<\lambda>} \oplus \ldots \oplus V^{<\lambda>}$ ($f_{[\lambda]}$ terms) is a maximal eigenspace of the Casimir operator S^2, and belongs to eigenvalue S(S+1). So this space may be labelled alternatively by the spin quantum number S (sec. I.4.7).
- Since $L^2(s)$ is 2-dimensional, only irreps with not more than two rows are afforded by this Hilbert space (sec. I.5.5).
- Each weight (uniquely characterized by an eigenvalue of S_z) occurs once in $V^{<\lambda>}$, for only one standard Weyl-Rumer or Gelfand tensor can be associated with a certain weight (sec. I.5.5 and I.5.6).

It remains to associate the spin quantum number S defined by:

$$S^2 O^{<\lambda>}(s_1, \ldots, s_N) = S(S+1) O^{<\lambda>}(s_1, \ldots, s_N)$$

with the corresponding partition $<\lambda>$.

From angular momentum theory [ref. 17, sec. 2.3] it is known that an eigenspace of S^2 with quantum number S is (2S+1)-dimensional. A subspace $V^{<\lambda>}$ with Young diagram:



is (p+1)-dimensional. (Apply to prove this formula (14) given in ref. 51). Comparison of dimensions shows that p = 2S, so S corresponds to $<\frac{1}{2}N + S, \frac{1}{2}N - S>$.

Notes.

1. The irreducibility postulate states: "Diagonalization of a spin-free Hamilton operator $H = H(\vec{r}) \otimes 1(\vec{s})$ induces a decomposition of the spin part of the Hilbert space into irreducible SU(2)-carrier spaces. Each SU(2)-carrier space belongs to one eigenvalue of H." (Therefore the dimension of such carrier spaces is called the "spin degeneracy" of the state). This group theoretical statement is equivalent to the quantum mechanical assertion: "The eigenstates of a spin-free Hamiltonian H are simultaneously eigenstates of the constants of motion S^2 and S_z. The spin-degenerate states are connected via ladder operators."
2. At this point it may seem surprising that the spin Hilbert space $L^2(s)^{\otimes N}$ will be decomposed by diagonalization of H, for H has the unit operator as its spin part. Subsequently we will see that the adaptation of $\mathcal{H}_M$ to S_N is responsible for this effect.
3. It is transparent how to generalize the theory to arbitrary spin σ. The Hilbert space is in general (2σ+1)-dimensional, the spin symmetry group is in general SU(2σ+1) and the maximum number of rows allowed in the Young diagram is then (2σ+1).

The next step in the adaptation of the model Hilbert space

$$\mathcal{H}_M \equiv (V_n^{\otimes N}) \otimes (L^2(s)^{\otimes N})$$

to the symmetry group G_H is the projection onto the antisymmetric subspace of $\mathcal{H}_M$. Doing this we will obviously profit by $L^2(s)^{\otimes N}$ already being

decomposed under SU(2), and hence under $S_N^{(s)}$.

At this point the choice must be made which of the possible bases of $L^2(s)^{\otimes N}$ is to be considered. Because the actual computations, reported on in the second half of this thesis, were performed on a basis of Weyl-Rumer tensors, we proceed in that representation. (See e.g. ref. 66 or ref. 15 for the corresponding derivations in terms of Gelfand tensors).

As we will see, the NP-structure matrix $\underline{X}^{[\lambda]}$ (defined in sec. I.5.1) will arise in the derivation below. It seems that the relevance of this matrix for the antisymmetrization problem has not been recognized before.

In the following theorem the operators on $\mathcal{H}_M$ are all factorized into two factors, the first acting on $V_n^{\otimes N}$ and the second on $L^2(s)^{\otimes N}$. The antisymmetrizer A reads in this notation:

$$A = 1/N! \sum_{P \in S_N} \zeta_P \, P \otimes P.$$

Theorem.

$$A(1 \otimes Y^{[\lambda]}_{k\ell}) = 1/N! \sum_{r=1}^{f^{[\lambda]}} \tau^{[\lambda]}_{kr} \, Y^{[\tilde{\lambda}]}_{\tilde{r}\tilde{k}} \otimes g^{[\lambda]}_{r\ell},$$

where:

$Y^{[\tilde{\lambda}]}_{\tilde{r}\tilde{k}}$ is the Young unit belonging to the standard tableaux $T^{[\tilde{\lambda}]}_{\tilde{r}}$ and $T^{[\tilde{\lambda}]}_{\tilde{k}}$, the associated tableaux of $T^{[\lambda]}_{r}$ and $T^{[\lambda]}_{k}$, respectively.

$\tau^{[\lambda]}_{kr}$ is the parity of $\sigma^{[\lambda]}_{kr}$, the permutation which converts $T^{[\lambda]}_{r}$ into $T^{[\lambda]}_{k}$.

$g^{[\lambda]}_{r\ell} \equiv \sum_j \eta^{[\lambda]}_{rj} \, Y^{[\lambda]}_{j\ell}$ (a natural unit).

$\eta^{[\lambda]}_{rj}$ is the (r,j)-element of the matrix $(\underline{X}^{[\lambda]})^{-1}$ defined in sec. I.5.1.

Proof.

$$A(1 \otimes Y^{[\lambda]}_{k\ell}) = 1/N! \sum_{P \in S_N} \zeta_P \, P \otimes (P Y^{[\lambda]}_{k\ell}) = 1/N! \sum_{P \in S_N} \zeta_P \, P \otimes \Big(\sum_j L^{D^{[\lambda]}}(P)_{jk} \, Y^{[\lambda]}_{j\ell}\Big).$$

Recall (sec. I.5.1) that

$$L\underline{D}^{[\lambda]}(P)^T = \underline{U}^{[\lambda]}(P^{-1})\,(\underline{X}^{[\lambda]})^{-1}$$

where $U^{[\lambda]}(P^{-1})_{k\ell}$ is the coefficient of P^{-1} in $Y^{[\lambda]}_{k\ell}$. It can be seen from the definition of $\underline{U}^{[\lambda]}(P)$ that:

$$\zeta_P \, U^{[\lambda]}(P^{-1})_{kr} = \tau^{[\lambda]}_{kr} U^{[\tilde{\lambda}]}(P)_{\tilde{r}\tilde{k}}$$

and so:

$$A(1 \otimes Y^{[\lambda]}_{k\ell}) = 1/N! \sum_{r=1}^{f_{[\lambda]}} \tau^{[\lambda]}_{kr} \Big(\sum_{P \in S_N} U^{[\tilde{\lambda}]}(P)_{\tilde{r}\tilde{k}} P \Big) \otimes \Big(\sum_{j=1}^{f_{[\lambda]}} \eta^{[\lambda]}_{rj} Y^{[\lambda]}_{j\ell} \Big) =$$

$$= 1/N! \sum_{r=1}^{f_{[\lambda]}} \tau^{[\lambda]}_{kr} Y^{[\tilde{\lambda}]}_{\tilde{r}\tilde{k}} \otimes \Big(\sum_{j=1}^{f_{[\lambda]}} \eta^{[\lambda]}_{rj} Y^{[\lambda]}_{j\ell} \Big).$$

Notes.

1. This theorem shows that the adaptation of spin space to SU(2) by the Young unit $Y^{[\lambda]}_{k\ell}$, followed by the antisymmetrization of $\mathcal{H}_M$, can also be performed in the following manner:

   (i) Adapt $V_n \otimes^N$ to $S_N^{(r)}$ by $Y^{[\tilde{\lambda}]}_{\tilde{r}\tilde{k}}$, $\tilde{r} = 1, \ldots, f_{[\tilde{\lambda}]}$.

   (ii) Adapt $L^2(s) \otimes^N$ to SU(2), using the natural units $g^{[\lambda]}_{r\ell}$. This gives also adaptation to $S_N^{(s)}$.

   (iii) Form linear combinations of the tensors so constructed. This amounts to the construction of the Clebsch-Gordon series $[\tilde{\lambda}] \otimes [\lambda] \downarrow S_N$ associated with the restriction of $S_N^{(r)} \otimes S_N^{(s)}$ to S_N.

   So, adaptation of the spin Hilbert space to SU(2) imposes, in cooperation with antisymmetrization, $S_N^{(r)}$-symmetry on the spatial part of the wavefunction. See e.g. ref. 63 for a discussion of some of the physical consequences of this observation.

2. In the preceding section it has been shown that diagonalization of $H_M(\vec{r})$ on $V_n \otimes^N$ decomposes this space into irreducible GL(n)-carrier spaces and hence, in the absence of spatial symmetry, to $S_N^{(r)}$. Using:

$$A(\tau^{[\lambda]}_{k\ell} Y^{[\tilde{\lambda}]}_{\tilde{\ell}\tilde{k}} \otimes 1) = A(1 \otimes Y^{[\lambda]}_{k\ell})$$

   and the theorem above it follows that the eigenstates of $H_M \equiv H_M(\vec{r}) \otimes 1(\vec{s})$ are adapted to $S_N^{(s)}$ and SU(2)-symmetry, i.e. they are eigenfunctions of S^2. So, antisymmetrization ensures that SU(2)-symmetry does not violate the irreducibility postulate. In more general terms it is the presence of S_N in the definition of G_H that takes care that the irreducibility postulate is not violated, even in the case of a spin-free Hamiltonian.

3. Since irreps $[\lambda]$ of $S_N^{(s)}$ with more than $(2\sigma+1)$ rows are not afforded by a general spin Hilbert space $L^2(s) \otimes^N$, it follows from the theorem above that the only physically realizable states of spin σ fermions are those with not more than $(2\sigma+1)$ columns in their Young diagrams.

6 5 A note on the calculation of matrix elements

The main purpose of this review: the presentation of a self-contained account of the group theoretical properties of N-particle systems has at this point been fulfilled. The second stage in a consistent and logical treatment of such systems would be the calculation of matrix elements, employing to this end all the simplifications that group theory can offer.

A discussion in the manner of the preceding chapters would certainly require more pages than the foregoing exposition. Also, this problem is still the subject of much current research, and therefore it would go far beyond the scope of this thesis to give an account including such exciting new developments as for example the use of double cosets in S_N or U(n)-shift operators for the evaluation of matrix elements.

In this section we will just briefly mention some of the existing methods.

If one wishes to employ the Gelfand basis a possible approach is from the side of $\Lambda(GL(n))$, writing the model Hamiltonian in terms of the generators of this Lie algebra (sec. I.6.3). One then calculates the matrix elements of the generators in this basis [34, 35]. Alternatively, one can use permutation group techniques, reducing the group theoretical part of the problem to the calculation of spin-free fractional parentage coefficients. This has been worked out in detail in ref. 67. Both methods ultimately end up with having to calculate 3n-j symbols of SU(2). (See also ref. 68 for a discussion of the permutation group method).

Using the Weyl-Rumer basis, consisting of canonical valence bond functions, which in our terminology are called "NP-projected tensors characterized by standard index tableaux" (sec. I.5.5), one can proceed in two different ways: "spin-free" or "conventional". To explain the difference we need the following theorem.

Theorem.

$$\langle A(\phi' \otimes Y^{[\lambda]}_{jg} O) \; |H| \; A(\phi \otimes Y^{[\lambda]}_{ig} O)\rangle = C \, \tau^{[\lambda]}_{ij} \, \langle Y^{[\tilde{\lambda}]}_{a\check{j}} \phi' \; |H| \; Y^{[\tilde{\lambda}]}_{a\check{i}} \phi \rangle ,$$

where: a, $1 \leq a \leq f_{[\lambda]}$, is arbitrary;

C is a constant which vanishes after normalization;

$\tau^{[\lambda]}_{ij}$ is the parity of $\sigma^{[\lambda]}_{ij}$;

$\phi, \phi' \in L^2(R^3)^{\otimes N}$;

$0 \in L^2(s)^{\otimes N}$ is chosen such that $Y^{[\lambda]}_{jg}0$ and $Y^{[\lambda]}_{ig}0$ are represented by a standard index tableau.

Proof. (Suppress $[\lambda]$, put a tilde over the operator to denote association). First we show.

$$Y^{\dagger}_{k\ell}Y_{pq} = c_{kp}Y^{\dagger}_{a\ell}Y_{aq},$$

where c_{kp} is a scalar independent of ℓ and q and a is arbitrary. Write:

$$Y^{\dagger}_{k\ell}Y_{pq} = P_{\ell}N_{\ell}\sigma_{\ell k}\sigma_{pq}N_{q}P_{q} \quad \text{(definition)}$$

$$= c'_{kp}P_{\ell}N_{\ell}\sigma_{\ell q}P_{q} \quad \text{(Von Neumann's theorem [36, p. 18])}$$

$$= (c'_{kp}/{}^{O}C_{\ell})\ P_{\ell}N_{\ell}N_{\ell}\sigma_{\ell q}P_{q} \quad \text{(because } N^2_{\ell} = {}^{O}C_{\ell}N_{\ell})$$

$$= c_{kp}P_{\ell}N_{\ell}\sigma_{\ell a}\sigma_{a\ell}\sigma_{\ell q}N_{q}P_{q} \quad (\sigma_{\ell a}\sigma_{a\ell} = (1),\ c_{kp} = c'_{kp}/{}^{O}C_{\ell})$$

$$= c_{kp}Y^{\dagger}_{a\ell}Y_{aq}$$

Using the theorem of the foregoing section, the matrix element becomes:

$$(\tfrac{1}{N!})^2 \sum_{k,\ell}\ \sum_{p,q} \tau_{kj}\tau_{pi}\eta_{k\ell}\eta_{pq}\,\langle \tilde{Y}_{\tilde{k}\tilde{j}}\phi' \,|H|\, \tilde{Y}_{\tilde{p}\tilde{i}}\phi \rangle \langle Y_{\ell g}0 \mid Y_{qg}0 \rangle$$

Using:

$$\tau_{kj}\tau_{pi} = (\tau_{kb}\tau_{bj})(\tau_{pb}\tau_{bi}) = (\tau_{kb}\tau_{bp})(\tau_{ib}\tau_{bj}) = \tau_{kp}\tau_{ij}$$

and:

$$\tilde{Y}^{+}_{\tilde{k}\tilde{j}}\tilde{Y}_{\tilde{p}\tilde{i}} = c_{\tilde{k}\tilde{p}}\tilde{Y}^{\dagger}_{\tilde{a}\tilde{j}}\tilde{Y}_{\tilde{a}\tilde{i}}$$

we get the matrix element:

$$\tau_{ij}\langle \tilde{Y}_{\tilde{a}\tilde{j}}\phi' \,|H|\, \tilde{Y}_{\tilde{a}\tilde{i}}\phi \rangle [\,(\tfrac{1}{N!})^2 \sum_{k,\ell}\ \sum_{p,q} \tau_{kp}c_{\tilde{k}\tilde{p}}\eta_{k\ell}\eta_{pq}\langle Y_{\ell g}0 \mid Y_{qg}0 \rangle]$$

The expression between square brackets is independent of a, $\tilde{i}$ and $\tilde{j}$, call it C and the theorem has been proved.

Notes.

1. This theorem gives a relation for all matrix elements which do not vanish because of symmetry reasons, for recall:
   - $Y^{[\lambda]}_{jg}C$ and $Y^{[\lambda]}_{ig}0$ are of definite weight (i.e. are eigenfunctions of S_z).
   - Each weight occurs only once in $\langle\lambda\rangle$.
   - Vectors of different weight are orthogonal.
   - Vectors belonging to different $\langle\lambda\rangle$ are orthogonal.
   - The spin part of H is the unit operator.

Hence all other matrix elements of H are zero.

2. The spin-dependent eigenvalue problem gives the same result as the spin-free eigenvalue problem defined in the space:

$$V_a^{<\tilde{\lambda}>} \equiv Y_{aa}^{[\tilde{\lambda}]} \; (V_n{}^{\otimes N})$$

for arbitrary a, $1 \leq a \leq f_{[\lambda]}$. The basis of this space consists of Weyl-Rumer tensors of the NP-type (sec. I.5.5).

The "conventional" way of calculating matrix elements proceeds by evaluating the antisymmetrizers in the left hand side of the expression in the theorem above. This yields:

$$\sum_{P \in S_N} \zeta_P \, \langle \Phi' \, | HP | \, \Phi \rangle \langle Y_{jg}^{[\lambda]} 0 \, | P | \, Y_{ig}^{[\lambda]} 0 \rangle .$$

Pauling [69] was the first to consider the evaluation of the spin matrix element arising in this evaluation. Or more precisely, he considered the special case of eigenfunctions of S^2 which are of the form:

$$(\alpha\beta - \beta\alpha)(\alpha\beta - \beta\alpha) \; \ldots \; (\alpha\beta - \beta\alpha).$$

These functions are obtained by the action of the Young unit belonging to the tableau:

1	3	5	7	·····
2	4	6	·····	

onto the function:

$$(\alpha\beta)(\alpha\beta) \; \ldots \; (\alpha\beta).$$

Clearly these projected functions have the eigenvalue $S(S+1) = 0$. Pauling imposed the Rumer diagram of the bra on the Rumer diagram of the permuted ket. This gives a superposition pattern, the analysis of which enables the formulation of very simple rules for the calculation of the spin matrix elements. Pauling's method has been extended to general spin multiplicities by, among others, Reeves [38], Cooper and McWeeny [39] and Sutcliffe [40]. Reeves published his results in the form of a computer algorithm which forms an important part of our programs. The reader is referred to the work of Van Berkel [47], who, following the Cooper and McWeeny paper, has described in detail the calculation of the spin matrix elements and the manner in which Reeves' algorithm has been implemented in our programs.

The "spin-free" method on the other hand starts working on the right hand side of the expression in the theorem above. This yields:

$$\langle \phi' | \; H(Y^{[\tilde{\lambda}]}_{\tilde{a}j})^{\dagger}(Y^{[\tilde{\lambda}]}_{\tilde{a}i}) \; | \phi \rangle \propto \langle \Phi' | \; HP^{[\tilde{\lambda}]}_{\tilde{j}} N^{[\tilde{\lambda}]}_{\tilde{j}} P^{[\tilde{\lambda}]}_{\tilde{j}} \sigma^{[\tilde{\lambda}]}_{\tilde{j}\tilde{i}} \; | \Phi \rangle \equiv \sum_{P \in S_N} [P]^{[\tilde{\lambda}]}_{\tilde{j}\tilde{i}} \langle \Phi' \; | HP | \; \Phi \rangle .$$

The scalars $[P]^{[\tilde{\lambda}]}_{\tilde{j}\tilde{i}}$ are sometimes called Pauling numbers [70], they clearly satisfy:

$$[P]^{[\tilde{\lambda}]}_{\tilde{j}\tilde{i}} \propto \zeta_P \langle Y^{[\lambda]}_{jg} 0 \; | P | \; Y^{[\lambda]}_{ig} 0 \rangle .$$

It is possible to evaluate Pauling numbers by purely group algebraic arguments, see Roël's dissertation [71]. Roël uses for the evaluation a double coset [68, 70] decomposition of S_N with respect to the normalizer [4, p. 15] of the row group $R^{[\tilde{\lambda}]}_{\tilde{j}}$. He then proves that all elements belonging to one double coset have the same Pauling number and that elements belonging to different double cosets have different Pauling numbers. These double cosets can be uniquely characterized by diagrams that are in one-to-one correspondence with Pauling's superposition patterns and their generalization to arbitrary spin multiplicities.

Notes.

1. All that has been said in this section about the calculation of matrix elements of the spin-free Hamiltonian holds for other spin-free operators too.
2. In the case of spin-dependent operators it is often advantageous to reduce the total Hilbert space first with respect to $S_N^{(r)}$ and $S_N^{(s)}$ and to apply the Wigner-Eckart theorem to space and spin matrix elements separately [72, 73]. However, not all spin-dependent operators are adapted to $S_N^{(r)}$ and $S_N^{(s)}$, although every physical operator is of course symmetric under S_N. In ref. 74 it has been discussed how operators can be classified according to irreps of $S_N^{(r)}$ and $S_N^{(s)}$ separately.

REFERENCES PART I

1. B.R. Judd,
   Operator Techniques in Atomic Spectroscopy.
   McGraw-Hill, New York, 1963.

2. B.G. Wybourne,
   Symmetry Principles and Atomic Spectroscopy.
   Wiley-Interscience, New York, 1970.

3. J. Tropfke,
   Geschichte der Elementar-Mathematik, Zw. Band,
   Dritte Auflage, Walter de Gruyter, Berlin, 1933.

4. L. Jansen, M. Boon,
   Theory of Finite Groups.
   North Holland, Amsterdam, 1967.

5. G.C. Shephard,
   Vector Spaces of Finite Dimension.
   Oliver & Boyd, Edinburgh, 1966.

6. H. Weyl,
   The Theory of Groups and Quantum Mechanics.
   Translated by H.P. Robertson.
   Reprinted by Dover, New York, 1950.

7. H. Weyl,
   The Classical Groups.
   Princeton University Press, 1946.

8. H. Boerner,
   Darstellungen von Gruppen.
   Zweite Auflage, Springer, Berlin, 1967.

9. C.W. Curtis, I. Reiner,
   Representation Theory of Finite Groups and Associative Algebras.
   Interscience, New York, 1962.

10. P.E.S. Wormer,
    Representatie Theorie van Eindige Groepen en Eindig Dimensionale
    K-algebra's.
    Report Technological University Delft, The Netherlands, May 1969.

11. D.J. Klein,
    Finite Groups and Semi-Simple Algebras in Quantum Mechanics.
    To appear in: Group Theory and its Applications III.
    Ed. E.M. Loebl, Academic Press.

12. W.H. Greub,
    Linear Algebra, 2nd edition.
    Springer, Berlin, 1963.

13. W. Miller Jr.,
    Symmetry Groups and their Applications.
    Academic Press, New York, 1972.

14. P.O. Löwdin,
Rev. Mod. Phys. 39, 259 (1967).

15. P.E.S. Wormer,
Different Orbitals for Different Spins.
Afstudeerverslag, Technological University Delft, The Netherlands,
June 1969.

16. M. Hausner, J.T. Schwartz,
Lie Groups; Lie Algebras.
Gordon and Breach, New York, 1968.

17. A.R. Edmonds,
Angular Momentum in Quantum Mechanics.
Princeton Univ. Press, Princeton, 1957.

18. F.D. Murnaghan,
The Theory of Group Representations.
Reprinted by Dover, New York, 1963.

19. F.D. Murnaghan,
The Unitary and Rotation Groups.
Spartan Books, Washington D.C., 1962.

20. W. Rudin,
Principles of Mathematical Analysis, 2nd edition.
McGraw-Hill, New York, 1964.

21. A. Messiah,
Quantum Mechanics.
North Holland, Amsterdam, 1965.

22. J.J. de Swart,
Group Theory, Part B.
Lecture Notes, Institute for Theoretical Physics,
University of Nijmegen, The Netherlands.

23. E.T. Whittaker,
Analytical Dynamics of Particles and Rigid Bodies, 4th edition.
Cambridge Univ. Press, Cambridge, 1937.

24. K.J. Lezuo,
J. Math. Phys. 13, 1389 (1972).

25. F.A. Matsen,
The Unitary Group Formulations of the Many-Electron Problem.
Preprint.

26. M. Hamermesh,
Group Theory.
Addison-Wesley, Reading, 1962.

27. D.E. Knuth,
The Art of Computer Programming, Vol. I.
Addison-Wesley, Reading, 1968.

28. S. Lang,
Algebra.
Addison-Wesley, Reading, 1965.

29. N. Jacobson,
Lie Algebras.
Interscience, New York, 1962.

30. L.C. Biedenharn,
J. Math. Phys. 4, 436 (1963).

31. J.D. Louck,
Am. J. Phys. 38, 3 (1970).

32. R.L. Hudson,
J. Math. Phys. 15, 1067 (1974).

33. P.A.M. Dirac,
The Principles of Quantum Mechanics, 4th edition.
Oxford Univ. Press, London, 1958.

34. J. Paldus,
J. Chem. Phys. 61, 5321 (1975).

35. J.F. Gouyet, R. Schranner, T.H. Seligman,
J. Phys. A 8, 285 (1975).

36. D.E. Rutherford,
Substitutional Analysis.
Edinburgh Univ. Press, Edinburgh, 1948.

37. F.A. Matsen, A.A. Cantu, R.D. Poshusta,
J. Phys. Chem. 70, 1558 (1966).

38. C.M. Reeves,
Comm. A.C.M. 9, 276 (1966).

39. I.L. Cooper, R. McWeeny,
J. Chem. Phys. 45, 226 (1966).

40. B.T. Sutcliffe,
J. Chem. Phys. 45, 235 (1966).

41. A.J. Coleman,
Adv. Quant. Chem. 4, 83 (1968).

42. F.A. Matsen,
Adv. Quant. Chem. 1, 59 (1964).

43. T. Yamanouchi,
Proc. Phys.-Math. Soc. Japan 19, 436 (1937).

44. P.E.S. Wormer, A. van der Avoird,
Int. J. Quant. Chem. 8, 715 (1974).
(Reprinted in this thesis).

45. G. Rumer, E. Teller, H. Weyl,
Gött. Nachr., Math.-Physik. Klasse, 499 (1932).

46. R. Pauncz,
Alternant Molecular Orbital Method.
Saunders, Philadelphia, 1967.

47. T. van Berkel,
Ab Initio Berekening van de Van der Waals Wisselwerking tussen
Twee He-Atomen.
Hoofdvakverslag, University of Nijmegen, 1974.

48. T.H. Seligman,
J. Math. Phys. 13, 876 (1972).

49. P.E.S. Wormer,
Chem. Phys. Lett. 5, 355 (1970).
(Reprinted in this thesis).

50. D.E. Littlewood,
The Theory of Group Characters, 2nd edition.
Oxford Univ. Press, London, 1950.

51. P.E.S. Wormer, A. van der Avoird,
J. Chem. Phys. 57, 2498 (1972).
(Reprinted in this thesis).

52. M. Brunet, T.H. Seligman,
Reports in Math. Phys. (1975)
(To be published).

53. W.J. Holman III, L.C. Biedenharn, in:
Group Theory and Its Applications.
Ed. E.M. Loebl.
Academic Press, New York, 1971.

54. M. Moshinsky,
Group Theory and the Many-Body Problem.
Gordon and Breach, New York, 1968.

55. E. Prugovecki,
Quantum Mechanics in Hilbert Space.
Academic Press, New York, 1971.

56. E.T. Copson,
Metric Spaces.
Cambridge Univ. Press, Cambrdige, 1968.

57. M. Boon,
Private communication.

58. V. Fock,
Z. Physik 98, 145 (1935).

59. H.V. McIntosh, in:
Group Theory and Its Applications II.
Ed. E.M. Loebl.
Academic Press, New York, 1971.

60. D.J. Klein,
Private communication.

61. P. Roman,
Advanced Quantum Theory.
Addison-Wesley, Reading, 1965.

62. W. Pauli,
Phys. Rev. 58, 716 (1940).

63. A. van der Avoird,
Identiteit.
Brakkenstein, Nijmegen, 1973.

64. E.C. Kemble,
The Fundamental Principles of Quantum Mechanics.
Reprinted by Dover, New York, 1958.

65. C.E. Wulfman, in:
Group Theory and Its Applications II.
Ed. E.M. Loebl.
Academic Press, New York, 1971.

66. A. van der Avoird,
Perturbation Theory for Intermolecular Forces.
Thesis, Eindhoven, 1968.

67. I.G. Kaplan,
Symmetry of Many-Electron Systems.
Academic Press, New York, 1975.

68. P. Kramer, T.H. Seligman,
Nucl. Phys. A 136, 545 (1969).

69. L. Pauling,
J. Chem. Phys. 1, 280 (1933).

70. D.J. Klein, C.H. Carlisle, F.A. Matsen,
Adv. Quant. Chem. 5, 219 (1970).

71. R.W.J. Roël,
Dissertation, forthcoming 1975/1976.
University of Amsterdam.

72. I.L. Cooper, J.I. Musher,
J. Chem. Phys. 57, 1333 (1972).

73. I.L. Cooper, J.I. Musher,
J. Chem. Phys. 59, 929 (1973).

74. A. van der Avoird, P.E.S. Wormer,
Chem. Phys. Lett. 14, 157 (1972).
(Reprinted in this thesis).

RUSSELL–SAUNDERS STATES ARISING FROM CONFIGURATIONS OF EQUIVALENT ELECTRONS *

P. E. S. WORMER
Paul M. Gross Chemical Laboratory, Duke University, Durham, North Carolina 27706, USA

Received 24 February 1970

A short method is given for the classification of the atomic Russell–Saunders coupling terms arising from an $(l)^n$ configuration consisting of n equivalent electrons. The method uses Young diagrams.

1. INTRODUCTION

The classification of the L–S coupling states of an atom or ion having more than two electrons outside a closed shell, is a cumbersome task if one uses the procedure of Slater [1]. An alternative and shorter route is presented which arrives at the same results.

In order to clarify the method a brief outline is presented to illustrate the connection between the irreducible representations of the general linear group GL(n) and the symmetric group S_k. Let V be an n-dimensional linear space and $V \otimes_k$ its k-fold (contravariant) tensorial product As is well known [2,3], we can characterize the irreducible GL(n)-subspaces of $V \otimes_k$ with the aid of Young diagrams This is so because all the partners in an irreducible representation of GL(n) can be generated by the action of a Young projector (ref [2], p 102) on the basis of $V \otimes_k$ Let the Young projector $Y_p^{(\lambda)}$ be constructed from the pth standard tableaux $\Theta_p^{(\lambda)}$ with shape $[\lambda]$ Let $Y_p^{(\lambda)}$ operate on the tensor $F_{i_1 i_2 \ldots i_k} \in V \otimes_k$, $1 \leq i_j \leq n$, $1 \leq j \leq k$. We can represent $Y_p^{(\lambda)} F_{i_1 i_2 \ldots i_k}$ graphically by the Young tableau that is obtained by replacing the digit j by the index i_j in $\Theta_p^{(\lambda)}$ By substituting in $\Theta_p^{(\lambda)}$ all the possible index sets $\{i_j \; 1 \leq j \leq k, \; 1 \leq i_j \leq n\}$ we generate a basis set for an irreducible GL(n)-subspace.

Of course, not all the $(n)^k$ tensors yield linearly independent basis vectors; only those basis vectors are linearly independent in which the graphical representations satisfy the following two conditions:

1-1 Going from left to right in a row the indices may not *in*crease;

1-2 Going down a column the indices must *de*crease.

The irreducible GL(n)-spaces generated in this way are irreducible under the subgroups U(n) and SU(n) of GL(n) as well An immediate consequence of these rules for eigenfunctions of the total spin angular momentum S^2 is that these functions can only be represented by Young diagrams of at most two rows As is easy to prove, the Young diagram $[\frac{1}{2}(k+s), \; \frac{1}{2}(k-s)]$ corresponds to a $(s+1)$-dimensional representation of SU(2)

2. L–S COUPLING CASE

Let us consider an atom with the electron configuration $(l)^k$. According to the Russell–Saunders rules we form from a product of k one-electron spin functions an eigenfunction $X(1, \ldots k)$ of S^2 with eigenvalue $\frac{1}{2}s(\frac{1}{2}s+1)$. This function can be represented by a Young tableau with pattern $[\frac{1}{2}(k+s), \; \frac{1}{2}(k-s)]$ Next we form a spatial product $\psi(1, \ldots k)$ of k orbitals with eigenvalue $l(l+1)$ of L^2, i.e. we construct an element of the k-fold tensorial product of the $(2l+1)$-dimensional carrier space of the irreducible representation $D^{(l)}$ of the full rotation group O(3). If one now antisymmetrizes the function $\psi(1, \ldots k) \otimes X(1, \ldots k)$ according to the Pauli principle, one finds [4] that the only non-vanishing components of ψ are those tensors that can be represented by Young tableaux with the shape $[2^{1/2(k-s)}, \; 1^s]$. Thus, the Pauli principle together with the fact that the atomic wavefunction ought to be an eigenfunction of S^2 causes the spatial part of the wavefunction

* Supported in part by NASA University Sustaining Grant NGR 34-001-005

to belong to specific irreducible representation spaces of SU($2l+1$) As pointed out in the introduction, it is very easy to write down the spatial wavefunctions that satisfy this symmetry restriction However, since the hamilton operator contains spatial electron-electron interaction terms, the total k-electron wavefunction ought to transform according to irreducible representations of the group O(3) $\subset$ SU($2l+1$) That means that we have to decompose the irreducible SU($2l+1$)-subspaces with respect to O(3) It is possible, using a formula derived by Goscinski and Ohrn [5], to find out which irreducible O(3)-subspaces are contained in an irreducible SU($2l+1$)-space Their formula is however difficult to apply because a knowledge of the characters of all the involved representations is required A faster way is the following write down the possible basis vectors of the irreducible SU($2l+1$) space that corresponds to a spin multiplicity $s+1$ That is, construct all the possible tensors in the Young tableau representation using the shape $[2^{\frac{1}{2}(k-s)}\ 1^s]$ taking into account the rules 1 1 and 1-2 Tabulate these tensors according to the quantum number M_L of the z-component of the angular momentum From this table it is immediately clear which representations are contained in the SU($2l+1$)-subspace

M_L \ L	5	4	3	3	2	1	1
5	2 2 1 0						
4	2 2 1 -1	2 1 1 0					
3	2 2 1 -2	2 2 0 1	2 1 1 1	2 0 1 0			
2	2 2 0 -2	2 1 1 -2	2 0 0 1	2 1 1 1	2 1 1 0		
1	2 2 -1 2	2 1 0 -2	2 0 1 2	2 0 0 -1	2 1 1 -1	2 -2 1 0	1 1 0 1
0	2 1 1 -2	2 0 0 2	2 1 1 2	2 1 0 -1	2 -2 1 -1	1 1 0 -2	1 0 0 -1

TABLE 1.

In order to make the argument concrete the triplet functions of a $(d)^4$ configuration are derived The orbital part should transform according to the Young tableau $[2, 1^2]$, so we get table 1. We see that a $(d)^4$ configuration gives rise to the following triplets

$$^3P\ \ ^3P\ \ ^3D\ \ ^3F\ \ ^3F\ \ ^3G\ \ ^3H$$

One can check the total multiplicity f of the orbital state by modifying Robinson's formula (ref [6] eq 3 282) to give

$$f \quad \left(\frac{s+1}{\frac{1}{2}(k+s)+1}\right)\left(\begin{matrix}2l+1\\ \frac{1}{2}(k+s)\end{matrix}\right)\left(\begin{matrix}2l+2\\ \frac{1}{2}(k-s)\end{matrix}\right)$$

In this example s 2, k 4 l - 2 hence f - (3 4)( ) - 45 This agrees with the sum of the multiplicities of L-states

$$f \quad 2\times 3 + 5 + 2\times 7 + 9 + 11 - 45$$

REFERENCES

[1] J C Slater Phys Rev 34 (1929) 1293
[2] H Börner Darstellungen von Gruppen second ed (Springer Berlin 1967)
[3] H Weyl The theory of groups and quantum mechanics (Dover Publications New York 1950)
[4] H S Johnston Rept Progr Phys 33 (1960) 66
[5] O Goscinski and Y Ohrn Intern J Quantum Chem 2 (1968) 845
[6] G de B Robinson Representation theory of the symmetric group (Edinburgh Univ Press 1961)

CLASSIFICATION OF INTERACTION OPERATORS WITH RESPECT TO MANY-PARTICLE PERMUTATION SYMMETRY

A VAN DER AVOIRD and P E S WORMER

Institute of Theoretical Chemistry, University of Nijmegen, The Netherlands

Received 6 March 1972

Two procedures are developed for the classification of interaction operators with respect to the permutation symmetry of a many (N) particle system which is a necessary first step for deriving selection rules for matrix elements of spin dependent operators over many particle wavefunctions. The first method based on character relations in the symmetric group S_N is applied to one- and two particle operators. The second, using Young diagrams is easily applied to the general case of n particle interaction operators.

1 Introduction

Recently Musher [1] has derived selection rules for matrix elements of spin dependent operators over many electron wavefunctions only using the property that such operators should be symmetric under permutations of the electrons whereas the wavefunctions must be antisymmetric. The formalism works for symmetric wavefunctions as well and, therefore can be applied to all systems containing N identical particles. It is based on the decomposition of the configuration space and the spin space separately into sub spaces which span the irreducible representations of the symmetric group S_N. A symmetric/antisymmetric wavefunction can be written as a sum of space functions which transform according to the irreducible representations of S_N, each term multiplied by a spin function which is a basis vector of the same/the associate representation [2, 3]. If the wavefunction is an eigenfunction of the total spin operator S^2 it corresponds to one specific irreducible representation. An analogous expansion can be made for the symmetric space - spin operators. After this decomposition of the wavefunctions and the operator matrix elements can be calculated by integrating over space and spin coordinates separately and applying the Wigner-Eckart theorem for the permutation group. The reduced matrix elements describe the dynamics of the problem; the 3-j symbols contain all information regarding the symmetry. Matrix elements are zero, by vanishing of all 3-j symbols, if the threefold (inner) product of the irreducible representations spanned by the components of the bra vector, the operator and the ket vector does not include the symmetric representation. Since the operator only contains some specific irreducible representations of S_N, this results in certain spin selection rules between the bra vector and the ket-vector [1].

Methods for calculating non-vanishing 3-j symbols of the permutation group are given by Gallup [4], Sullivan [5], Cooper and Musher [6]. In order to apply this procedure for the simplification of matrix elements explicitly one must find out which irreducible representations of S_N are carried by the separate space and spin operators that constitute the N particle interaction operator.

(1) One particle operators

$$\{f(i) | i = 1, \quad N\}$$

are contained in the general interaction operator

$$O_1 = \sum_{i=1}^{N} a(r_i)\, b(\sigma_i)$$

which describes e.g. spin-orbit coupling or the Fermi hyperfine interaction.

(ii) Two-particle operators

$$\{f(i\ j)|i=1,\quad ,N, j=1,\quad ,i\quad 1\}$$

occur in interactions of the type

$$O_2=\sum_{i=1}^{N}\sum_{j=1}^{i-1} g(r_i,r_j)h(\sigma_i,\sigma_j)$$

They can be symmetric, $f(i,j)=f(j\ i)$ such is electron electron, spin spin coupling, or antisymmetric $f(i,j)=-f(j,i)$, in case of vector forces [7] Non symmetric two-particle operators appearing in the spin other-orbit coupling [8], can be written as sums of a symmetric and an antisymmetric part An important class of two-particle spin operators is given by various effective hamiltonians, symmetric if they are of the Heisenberg type [9], antisymmetric in some more extended models [10]

(iii) Many-particle operators

$$\{f(i,j,k,\quad)|N\geqslant i>j>k>\quad \geqslant 1\}$$

involving interactions between more than two particles simultaneously, are not found in any "physical" hamiltonian Still they do arise in effective interaction operators [11, 12]

In this letter we first use the character relations in order to prove the decomposition of the spaces spanned by the one particle and two-particle operators given by Musher [1] and Gallup [4] We then treat the general case of n particle operators in the N particle Hilbert space

2 One and two-particle operators

(a) The one-particle operators

$$\{f(i)|i=1,\quad ,N\} \tag{1}$$

carry an N dimensional (reducible) representation Γ_1 of the symmetric group S_N which should be decomposed as a direct sum of irreducible representations The character of this "permutation representation Γ_1 for a certain class of S_N is derived by acting with a permutation from this class on the basis of 1 particle operators The character equals the number of basis vectors mapped on itself The class structure of S_N is completely determined by the cycle structure of the permutations, so that we can denote an arbitrary class of permutations consisting of k 1 cycles l 2 cycles m 3-cycles etc as

$$(k\ l,m\quad)=(1^k\ 2^l\ 3^m\quad), \tag{2}$$

with

$$k+2l+3m+\quad =N$$

If an operator from this class acts on the basis (1) all 2 cycles and larger cycles interchange operators Each 1 cycle leaves one operator invariant So the character of this permutation representation is

$$\chi^{\Gamma_1}_{(k\ l\ m\quad)}=k \tag{3}$$

This character must be written as a sum of irreducible characters of S_N which can be derived by an algorithm described in Hamermesh [13] Using the partition notation for the irreducible representations of S_N one finds that

$$\chi^{[N]}_{(k\ l\ m\quad)}=1\ , \tag{4}$$

$$\chi^{[N\quad 1\ 1]}_{(k\ l\ m\quad)}=k\quad 1 \tag{5}$$

so that the unique decomposition of the representation Γ_1 is given by

$$\Gamma_1=[N]\quad [N\quad 1\ 1] \tag{6}$$

(b) The two particle operators

$$\{f(i\ j)|i=1\qquad N\ j=1\qquad i\quad 1\} \tag{7}$$

span a $\frac{1}{2}N(N\quad 1)$ dimensional representation Let us again act with a permutation operator of the arbitrary class $(1^k\ 2^l\ 3^m,\quad)$ on all $f(i\ j)$ The k 1 element partitions leave a certain $f(i\ j)$ invariant if they include both (i) and (j) A sequence of k numbers contains $\frac{1}{2}k(k\quad 1)$ different pairs $(i\ j)$ with $i>j$ so that the 1 element partitions in the permutation operator leave $\frac{1}{2}k(k\quad 1)$ operators $f(i\ j)$ unaltered Besides if we assume that the operators $f(i\ j)$ are all symmetric or all antisymmetric under the transposition (ij) an $f(i\ j)$ is left invariant or turned into $\quad f(i\ j)$ respectively by this transposition Therefore every 2 cycle in the permutation operator maps one $f(i\ j)$ on itself in the symmetric case or on its negative for antisymmetric operators Since all larger cycles necessarily interchange the $f(i\ j)$ the characters of the representations spanned by the symmetric and antisymmetric 2 particle operators are respectively

$$\chi^{\Gamma_{2\ sym}}_{(k\ l\ m\quad)}=\tfrac{1}{2}k(k\quad 1)+l \tag{8}$$

$$\chi^{\Gamma_{2\,\mathrm{anti}}}_{(k,l,m,\ldots)} = \tfrac{1}{2}k(k-1) - l\,. \tag{9}$$

Inspecting the irreducible characters of S_N given by Hamermesh [13].

$$\chi^{[N-2,2]}_{(k,l,m,\ldots)} = \tfrac{1}{2}(k-1)(k-2) + l - 1\,, \tag{10}$$

$$\chi^{[N-2,1^2]}_{(k,l,m,\ldots)} = \tfrac{1}{2}(k-1)(k-2) - l\,, \tag{11}$$

we find that

$$\Gamma_{2\,\mathrm{sym}} = [N] \oplus [N-1,1] \oplus [N-2,2]\,, \tag{12}$$

$$\Gamma_{2\,\mathrm{anti}} = [N-1,1] \oplus [N-2,1^2]\,. \tag{13}$$

Although possible in principle, the treatment of three- or more-particle operators in a similar manner would become very laborious. For this reason we shall consider the general case of n-particle operators by an alternative technique, which basically is derived from character relations as well [14].

3. General theory

The general problem is to find, which representations are carried by the set of all n-particle operators in N-particle space ($1 \leqslant n \leqslant N$). Without loss of generality the n-particle operators are assumed to be symmetry adapted with respect to the permutation group S_n, they are irreducible tensor operators which transform according to the representation $[\lambda]$ of S_n (i.e., the symmetric or the antisymmetric one in case of 2-particle operators). If this would not be the case for any practical set of operators we could always expand them as linear combinations of such symmetry adapted operators. We denote these operators by

$$\{f^{[\lambda]}_k(I_n) | k = 1, \ldots, f_\lambda\}\,, \tag{14}$$

where I_n stands for a set of n different particle coordinates (taken from the total set I_N of N particle coordinates) and f_λ is the dimension of the irreducible representation $[\lambda]$. Special cases of this formula are given by the expressions (1) and (7). We factorize the N-particle Hilbert space as a twofold tensorial product, one factor being the space of functions on I_n, the other one the space of functions on the set of coordinates which do not occur in I_n. This difference set is denoted by $I_N \backslash I_n$. This factorization of the N-particle Hilbert space corresponds to an equivalent factorizing of the operators on this space. So we write the operators $f^{[\lambda]}_k(I_n)$ in N-particle space as:

$$f^{[\lambda]}_k(I_n) \otimes 1(I_N \backslash I_n)\,, \tag{15}$$

where $1(I_N \backslash I_n)$ stands for the identity operator in the space of the particles in the difference set $I_N \backslash I_n$.

Consider a particular one of these operators, for instance:

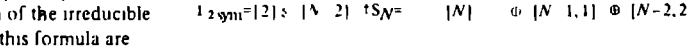
$$f^{[\lambda]}_k(1, \ldots, n) \otimes 1(n+1, \ldots, N)\,. \tag{16}$$

It is apparent that $1(n+1, \ldots, N)$ transforms according to the symmetric representation of S_{N-n}, corresponding to the partition $[N-n]$. Since $f^{[\lambda]}_k$ transforms according to the representation $[\lambda]$ of S_n, the operator given by (16) must carry the (irreducible) representation $[\lambda] \otimes [N-n]$ of the productgroup $S_n \otimes S_{N-n}$, which is a subgroup of S_N. The total set of n-particle operators in N-particle space can be generated from this particular operator (16) by the coset generators of $S_n \otimes S_{N-n}$ in S_N. The representation spanned by these operators is thus obtained by inducing the representation $[\lambda] \otimes [N-n]$ to the permutation group S_N. It is symbolically denoted by $[\lambda] \otimes [N-n] \uparrow S_N$. The rules for decomposing such representations which have been given by Littlewood [14], are as follows After writing the Young diagram for the irreducible representation $[\lambda]$, one adds the boxes of the (single row) diagram $[N-n]$ one by one to the boundary of $[\lambda]$ so that

(i) the augmented diagram remains a proper Young diagram,

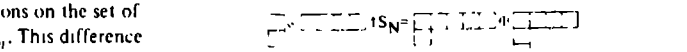
(ii) not more than one box is added to each column of $[\lambda]$

For example, let us treat the two-electron operators again ($n = 2$). If they are symmetric, $[\lambda] = [2]$ and the decomposition reads.



$\Gamma_{2\,\mathrm{sym}} = [2] \otimes [N-2] \uparrow S_N = [N] \oplus [N-1,1] \oplus [N-2,2]$

For antisymmetric operators with $[\lambda] = [1^2]$, we find.



$\Gamma_{2\,\mathrm{anti}} = [1^2] \otimes [N-2] \uparrow S_N = [N-1,1] \oplus [N-2,1^2]$

4 Conclusion

We have derived two methods for decomposing the representation spanned by the n particle operators in N particle Hilbert space into irreducible representations of S_N. The second method using Young diagrams is much easier to handle than the first one based on character relations. Since this decomposition can be applied to both the space and spin operators separately general spin dependent operators can be decomposed and their matrix elements simplified by the techniques described in the introduction.

Acknowledgement

We like to thank Professor J.I. Musher for suggesting this problem and for valuable discussions. One of us (AvdA) is also grateful to him for a stimulating stay at Yeshiva University (New York). This research was partially supported by the National Science Foundation and by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) via the Netherlands Foundation for Chemical Research (S.O.N.).

References

[1] J.I. Musher, Chem. Phys. Letters 11 (1971) 153.
[2] H. Weyl, The theory of groups and quantum mechanics (Dover Publications, New York, 1950) p. 370 ff.
[3] M. Kotani, A. Amemiya, T. Ishiguro and T. Kimura, Table of molecular integrals (Maruzen, Tokyo, 1955).
[4] G.A. Gallup, J. Chem. Phys. 52 (1970) 893.
[5] J.J. Sullivan, Phys. Rev. A5 (1972) 29.
[6] I.L. Cooper and J.I. Musher, J. Chem. Phys., to be published.
[7] A. de-Shalit and I. Talmi, Nuclear shell theory (Academic Press, New York, 1963).
[8] M. Blume and R.E. Watson, Proc. Roy. Soc. A270 (1962) 127.
[9] P.W. Anderson, Solid State Phys. 14 (1963) 99.
[10] T. Moriya, Phys. Rev. 120 (1960) 91.
[11] B.R. Judd, Advan. Chem. Phys. 14 (1969) 91.
[12] E.A. Matsen, D.J. Klein and D.C. Doyle, J. Phys. Chem. 75 (1971) 1866.
[13] M. Hamermesh, Group theory and its applications to physical problems (Addison-Wesley, Reading, 1962) p. 206.
[14] D.E. Littlewood, The theory of group characters, 2nd Ed. (Oxford Univ. Press, London, 1950) p. 92.

Transformation Properties of Antisymmetric Spin Eigenfunctions under Linear Mixing of the Orbitals

PAUL E S WORMER AND AD VAN DER AVOIRD
Institute of Theoretical Chemistry, University of Nijmegen, The Netherlands
(Received 8 May 1972)

After recalling the duality between the general linear group $GL(m)$, represented by its N-fold inner product, and the permutation group S_N, we have given a survey of its quantum chemical consequences. It causes the one-to-one correspondence between the total spin quantum number and the permutation symmetry of N-electron spin functions, and, via the Pauli principle which imposes permutation symmetry on the spatial part also, it leads to specific properties of antisymmetric spin eigenfunctions under orbital transformations. Such functions can be classified according to the irreducible representations of $GL(m)$. For special orbital transformations, often occurring in quantum chemistry, which mix only orbitals in different subsets among each other, we have derived how the transformation of the N-electron wavefunctions simplifies, by a reduction of the representations of $GL(m)$. The theory is illustrated by an example and some applications are discussed.

I. INTRODUCTION

The last years have shown among theoretical chemists a renewed interest in the applications of the permutation group to many-body problems. Kotani *et al.*[1] and Johnston[2] revived the early work of Hund, London, Heitler, Wigner, and, especially, Weyl.[3] The theory was worked out further by Matsen and co-workers[4] in a series on "spin-free quantum chemistry". It was applied to variational calculations by Goddard,[5] Harris,[6] and Morrison and Gallup,[7] whereas van der Avoird,[8] Amos and Musher[9] and Klein[10] continued the work of Eisenschitz and London[11] on permutation symmetry adapted perturbation theory.

One aspect of the theory is missing in this recent work: The duality between the representations of the permutation group S_N on the one hand and the general linear group in m dimensions $GL(m)$ on the other. These groups are related via their representations carried by Nth rank tensor space. Still, this relationship is of paramount importance. In the first place it forms the basis of the connection between the spin quantum number and permutation symmetry: N-electron eigenfunctions of S^2, which carry irreducible representations of $GL(2)$ and its subgroup $SU(2)$, must span certain irreducible representations of the permutation group S_N. Secondly, the Pauli principle, requiring the antisymmetry of the total many-electron wavefunction, imposes the permutation symmetry on the spatial part as well. Using the duality between S_N and $GL(m)$ again, it follows that the spatial function must have specific behavior under orbital transformations, which can be studied by considering the group $GL(m)$.

The global representation theory of $GL(m)$ on which this paper is based, originates from Schur and has been outlined in great detail by Weyl,[3] who was the first to recognize its quantum mechanical importance. The books by Boerner[12] and Murnaghan[13] give a good survey of the mathematical background.

In many practical cases one is not interested in general orbital transformations, but, given a partitioning of the orbitals in different subsets, in those transformations which mix only orbitals belonging to the same subset. Examples are given by:

the hybridization of orbitals on the separate atoms in a molecular Valence Bond calculation,

orthogonalization of orbitals in different sets by intraset transformations (According to the pairing theorem such transformations can orthogonalize an orbital in a definite set to all orbitals except one in another set.),

mixing schemes in pair-correlated DODS methods, such as AMO [14] or extended VB,[15]

Roothaan's procedure[16] of simplifying the open-shell Hartree–Fock problem by transformation of the closed and open shell orbitals separately.

In this paper we have derived which antisymmetrized eigenfunctions of S^2 are mixed under such "partitioned" orbital transformations and indicated a way to calculate the mixing coefficients. The theory is outlined for a system consisting of two parts, but, by induction, is easily extensible to many subsystems. It could be considered as a supplement to Matsen and Klein's "aggregate theory,"[4] regarding the transformation properties of aggregate states. Before we proceed to our results we shall give a brief survey of the general formalism which is extensively, but rather untransparently, described by Weyl.[3]

II. REVIEW OF THE DUALITY BETWEEN S_N AND $GL(m)$

For the construction of the wavefunction, let us begin with an m-dimensional orbital set spanning a linear

space V_m^r,

$$V_m^r: \quad \{\phi_i(r) \mid i = 1, \cdots, m\} \tag{1}$$

The general linear group $GL(m)$ consists of all non singular linear mappings (orbital transformations) γ,

$$\gamma: \quad V_m^r \rightarrow V_m^r \tag{2}$$

Formation of all N electron product functions amounts to constructing a tensorial product space spanned by Nth rank (purely contravariant) tensors,

$$V_m^r \otimes^N \quad \{\Phi_{i_1 i_2 \cdots i_N} \mid i_1, i_2, \cdots, i_N = 1, \cdots, m\}, \tag{3}$$

where

$$\Phi_{i_1 i_2 \cdots i_N} = \phi_{i_1}(r_1) \otimes \phi_{i_2}(r_2) \otimes \cdots \otimes \phi_{i_N}(r_N)$$

Similarly, the N-electron spin space $V_2^s \otimes^N$ is an N-fold tensorial product of 2 dimensional one electron spin spaces V_2^s

The linear space $V_m^r \otimes^N$ is stable under both the permutation group S_N and the N fold inner product group $[GL(m)]^N$ The latter consists of the tensor products of mappings $[\gamma]^N$,

$$[\gamma]^N \quad V_m^r \otimes^N \rightarrow V_m^r \otimes^N, \tag{4}$$

which are defined by

$$[\gamma]^N \Phi_{i_1 i_2 \cdots i_N} = \gamma\phi_{i_1}(r_1) \otimes \gamma\phi_{i_2}(r_2) \otimes \cdots \otimes \gamma\phi_{i_N}(r_N), \tag{5}$$

that is, by a simultaneous transformation of all orbitals in $\Phi_{i_1 i_2 \cdots i_N}$ under γ If $\mathbf{D}(\gamma)$ is the matrix of γ, then $[\mathbf{D}(\gamma)]^N$, the Nth power Kronecker product of $\mathbf{D}(\gamma)$, is the matrix of $[\gamma]^N$ The group consisting of these Kronecker product matrices is a faithful representation of $GL(m)$ (and of the isomorphic group $[GL(m)]^N$) carried by Nth rank tensor space This tensor space can be decomposed with respect to the permutation group S_N using the matrix element operators or Wigner operators (which generate minimal right ideals of the group algebra of S_N [12])

$$W_{ij}^{[\lambda]} = (f_{[\lambda]}/N!) \sum_{P \in S_N} \mathbf{D}_{ji}^{[\lambda]}(P^{-1}) P \tag{6}$$

Here $\mathbf{D}_{ji}^{[\lambda]}$ is a matrix element of the $f_{[\lambda]}$ dimensional irreducible matrix representation $[\lambda]$ of S_N (which can be denoted by a partition of N) Because the elements of S_N commute with all the elements of the group $[GL(m)]^N$, it follows that a complete reduction of $V_m^r \otimes^N$ under S_N brings along the following complete reduction under $[GL(m)]^N$ and, consequently, under $GL(m)$ [12]

$$V_m^r \otimes^N = \sum_{(\lambda)} \oplus R^{(\lambda)} \tag{7}$$

partitions of N

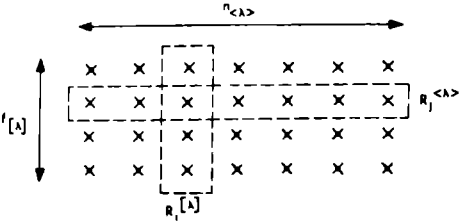


FIG 1 Simultaneous decomposition of the space $R^{(\lambda)}$ in basis vectors of the irreducible representations $[\lambda]$ of S_N and $\langle\lambda\rangle$ of $GL(m)$ These basis vectors can be obtained by applying the Wigner operators $W_{j1}^{[\lambda]}$ $j = 1, \cdots f_{[\lambda]}$, [Formula (6)] on different product configurations which can be constructed from the given orbital set

with

$$R^{(\lambda)} = \sum_{i=1}^{n_{\langle\lambda\rangle}} \oplus R_i^{[\lambda]} = \sum_{j=1}^{f_{[\lambda]}} \oplus R_j^{\langle\lambda\rangle}$$

$R^{[\lambda]}$ spans the $f_{[\lambda]}$ dimensional representation $[\lambda]$ of S_N which occurs $n_{\langle\lambda\rangle}$ times in $V_m^r \otimes^N$, whereas $R_j^{\langle\lambda\rangle}$ spans the $n_{\langle\lambda\rangle}$ dimensional representation $\langle\lambda\rangle$ of $GL(m)$ occurring $f_{[\lambda]}$ times Here Robinson's[18] notation is used Both dimensions, $f_{[\lambda]}$ and $n_{\langle\lambda\rangle}$, are easily calculated by means of hook graphs and m graphs, respectively [18 19] Schematically this decomposition (7) is shown in Fig 1.

Henceforth we assume that the different spaces $R_i^{[\lambda]}$, $i = 1, \cdots, n_{\langle\lambda\rangle}$, which span the irreducible representations $[\lambda]$ of S_N, carry identical matrix representations Weyl's proof shows that, in this case, the spaces $R_j^{\langle\lambda\rangle}$, $j = 1, \cdots, f_{[\lambda]}$, carry identical matrix representations $\mathbf{D}^{\langle\lambda\rangle}$ of $GL(m)$, so that we obtain the following decomposition of $[\mathbf{D}(\gamma)]^N$

$$[\mathbf{D}(\gamma)]^N = \sum_{(\lambda)} \oplus f_{[\lambda]} \mathbf{D}^{(\lambda)}(\gamma) \tag{8}$$

partitions of N

Since the elements of $\mathbf{D}^{(\lambda)}(\gamma)$ are integral functions (Nth order homogeneous polynomials) of the elements of $\mathbf{D}(\gamma)$, these irreducible representations of $GL(m)$ are called integral [12 13] The reason that this one to-one correspondence between the irreducible representations of S_N and $GL(m)$ has important consequences in the quantum mechanics of N electron systems is explained in the next two sections

III SPIN QUANTUM NUMBER AND PERMUTATION SYMMETRY

The ξ-component of the spin operator, S_ξ, for one electron is related to the unitary mappings in two-dimensional spin space V_2^s by

$$U_\xi(\phi) = \exp(i\phi S_\xi), \tag{9}$$

where $U_\xi(\phi)$ can be thought to represent a rotation around the ξ axis over angle ϕ. These "rotations" form the group $SU(2)$ of all two-dimensional unitary transformations with unit determinant. In case of N electrons the same formula (9) holds for the tensor product operators $[U_\xi(\phi)]^N$, constituting the group $[SU(2)]^N$ of mappings in $V_2{}^a \otimes^N$.

Using Formula (9) it is easily proved that the group $[SU(2)]^N$ commutes with the total spin operator S^2 and, invoking Schur's lemma, that N-electron spin functions which are basis vectors for the irreducible representations of $[SU(2)]^N$, and of $SU(2)$, are eigenfunctions of S^2. The duality between the representations of $[U(2)]^N$ and S_N and the property that irreducible representations of $U(2)$ stay irreducible under the subduction $U(2) \downarrow SU(2)$, then leads to the one-to-one correspondence between the irreducible representations of S_N and the total spin quantum number. Eigenfunctions of S^2 are basis vectors for the irreducible representations of S_N, corresponding to partitions of N, graphically represented by Young diagrams. The dimensionality 2 of the spin space causes at most two-element partitions (two-row diagrams) to occur, so that one can write a basis element of $V_2{}^a \otimes^N$ as

$$|[N/2+S, N/2-S], M_s, j\rangle, \qquad (10)$$

where S denotes the total spin quantum number, M_s the z component of the spin, and the index $j = 1, \cdots, f_{[N/2+S, N/2-S]}$.

IV. TRANSFORMATION PROPERTIES OF THE SPATIAL WAVEFUNCTIONS

If one neglects spin-dynamics, then S and M_s are good quantum numbers and the spin part of the N-electron wavefunction must be an eigenfunction of S^2. In order to construct the spatial wavefunction one must first select an appropriate configuration $|\Phi_{I_N}\rangle$, where I_N corresponds to a set of specific orbital indices $I_N = \{i_1, i_2, \cdots, i_N\}$. The total N-electron wavefunction, which must be antisymmetrized according to the Pauli principle, can then be written as[3]

$$|\psi_{I_N, k}{}^{S\ M_s}\rangle = \sum_{j=1}^{f_{[N/2+S\ N/2-S]}} |[N/2+S, N/2-S], M_s, j\rangle \otimes W_{jk}{}^{[2^{N/2-S}, 1^{2S}]}\, |\Phi_{I_N}\rangle. \qquad (11)$$

In this expression $[2^{N/2-S}, 1^{2S}]$ is the associate (two-column) representation of $[N/2+S, N/2-S]$. The index k of the Wigner operator has to be chosen such that the resulting tensor is not equal to zero. In general, several choices of k, resulting in tensors with different parentage, are possible.

Summarizing, it follows that spin symmetry together with the Pauli principle imposes definite permutation symmetry on the spatial part of the wavefunction. The latter symmetry in turn, because of the duality between S_N and $GL(m)$, causes a reduction of $V_m{}^a \otimes^N$ under $GL(m)$.

When neglecting spin terms in the Hamiltonian, a general Hamilton matrix element takes the form

$$\langle \psi_{I_{N'}, k'}{}^{S', M_{s'}} | H | \psi_{I_N, k}{}^{S\ M_s}\rangle = \delta_{S'S}\delta_{M_{s'}M_s} f_{[\lambda]} \times \langle \Phi_{I_{N'}} | H | W_{k'k}{}^{[\lambda]} \Phi_{I_N}\rangle, \qquad (12)$$

where $[\lambda]$ stands for $[2^{N/2-S}, 1^{2S}]$.

For arriving at this result we used the property that the Wigner operators commute with H and the relation

$$W_{j'k'}{}^{[\lambda']\dagger} W_{jk}{}^{[\lambda]} = \delta_{[\lambda'][\lambda]}\delta_{j'j} W_{k'k}{}^{[\lambda]}. \qquad (13)$$

Formulas (12) and (13) show in the first place that the carrier spaces $R_j{}^{(\lambda)}$, $j = 1, \cdots, f_{[\lambda]}$, of $GL(m)$ are non-interacting for different $\langle\lambda\rangle$ and different j. Secondly, instead of writing the matrix elements over antisymmetric space–spin functions $|\psi\rangle$, one can also write matrix elements over the space functions $W_{jk}{}^{[\lambda]}\, |\Phi_{I_N}\rangle$ only. The reason is that the matrix elements over such space functions are independent of j, and identical (except for multiplication by a constant) to the matrix elements over the space spin functions. Concluding, we can take just one of the carrier spaces $R_j{}^{(\lambda)}$ of $GL(m)$ in order to construct all matrix elements in the secular problem. Neglecting spin dynamics, we may replace the Pauli principle and the spin symmetry by the single postulate: A physically allowable N-electron spatial wavefunction must be expanded in partner basis elements of an integral representation of $GL(m)$, characterized by a Young diagram with an most two columns.

Although this postulate is equivalent to Matsen's,[20] which concerns the permutation group S_N, the formulation in terms of $GL(m)$ emphasizes directly the spatial transformation properties of the wavefunction. This postulate has one drawback. We change the degeneracy of the system from $2S+1$ to $f_{[2^{N/2-S}\ 1^{2S}]}$, where the latter degeneracy cannot be resolved by any physical means whatever (Ref. 3, p. 321).

Let us illustrate the quantum chemical application of this theory by an example. The O_2 molecule counts 10 atomic orbitals in the K and L shells that can be considered in a valence bond calculation, or the same number of molecular orbitals if we start with an MO treatment on that basis. Suppose we wish to construct an antisymmetric wavefunction for the triplet ground

state of O_2 containing 16 electrons This triplet function must be a basis element of the representation $\langle 2^7, 1^2\rangle$ of $GL(10)$ A simplification is obtained by noting that this representation has the same dimension as the representation $\langle 2^7, 1^2\rangle$ of $SU(10)$, which is equivalent to the "hole" representation $\langle 2, 1^2\rangle$ of $SL(10)$ [21] This hole representation is defined such that its Young diagram added to the original diagram as in Fig 2 yields the Young diagram for m doubly occupied orbitals

The dimension $n_{\langle 2\,1^2\rangle}$, which equals the number of triplet configurations that can be constructed from the given orbital basis, is easily calculated by Robinson's formula[18]

$$n_{\langle\lambda\rangle}=G_{\langle\lambda\rangle}(m)/H_{\langle\lambda\rangle}, \tag{14}$$

where $G_{\langle\lambda\rangle}(m)$ is the product of numbers in the m graph and $H_{\langle\lambda\rangle}$ is the product of hook lengths The dimension is 990

Thus means that an orbital transformation of the original 10 orbital basis set, for instance a Lowdin orthogonalization of the VB orbitals or a DODS mixing of the MO's, would mix 990 triplet states of O_2 If the applied orbital transformation $\mathbf{D}(\gamma)$ is predetermined, as for the Lowdin orthogonalization, the weights of all mixed configurations are fixed they are the matrix elements of $\mathbf{D}(\gamma)^{(2^{N/2-S}\,1^{2S})}$ In the case of the DODS procedure, the matrix elements of $\mathbf{D}(\gamma)$ are used as variation parameters The theory of this section then tells immediately that the DODS approach is equivalent to a full configuration interaction with somewhat restricted CI coefficients that are homogeneous polynomials of the DODS parameters

V. PARTITIONED ORBITAL TRANSFORMATIONS

So far we have outlined the basic theory derived by Schur and Weyl From this theory follows the manner in which an antisymmetric spin eigenfunction behaves under arbitrary orbital transformations In many practical cases, however (examples are given in the introduction), one can divide the orbitals in different subsets and restrict the transformations between orbitals of the same subset In these cases the matrix $\mathbf{D}(\gamma)$ has a blocked form We have derived explicitly how the transformation of the many electron wave

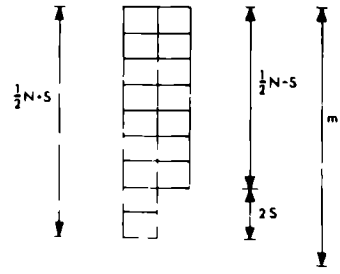


FIG 2 Young diagram for the triplet ($S=1$) state of O_2 number of electrons $N=16$, number of orbitals $m=10$

function simplifies under such orbital transformations The result is described in this section for a two-subset partitioning of the orbital set, but, by induction, can be easily extended to the general case

A division of the orbital set into two subsets is equivalent to the following decomposition of the orbital space

$$V_m=V_{m_1}\oplus V_{m_2} \tag{15}$$

We consider orbital transformations of the form

$$\gamma=\gamma_1\oplus\gamma_2, \tag{16}$$

where γ_1 is the restriction of γ to V_{m_1} and γ_2 to V_{m_2} The mappings γ of this form constitute a group which we denote by $GL(m_1+m_2)$ Since $m_1+m_2=m$, this group is clearly a subgroup of $GI(m)$ Thinking in terms of matrices instead of linear mappings $GL(m_1+m_2)$ can be defined as the group of matrices with the blocked structure

$$\mathbf{D}(\gamma)=\mathbf{D}(\gamma_1)\oplus\mathbf{D}(\gamma_2), \tag{17}$$

where $\mathbf{D}(\gamma_1)$ is a $m_1\times m_1$ matrix and $\mathbf{D}(\gamma_2)$ a $m_2\times m_2$ matrix

The problem we have to solve is to find the behavior of the irreducible representations $\langle\lambda\rangle$ of $GL(m)$ under the subduction $GI(m)\downarrow GI(m_1+m_2)$ First we decompose the λth rank tensor space $V_m\otimes^\lambda$ by a generalization of Newton's binomial theorem for noncommuting factors

$$V_m\otimes^\lambda=(V_{m_1}\oplus V_{m_2})\otimes^\lambda = \sum_{n=0}^{\lambda}\sum_{k=1}^{\binom{\lambda}{n}}\oplus C_k\,(V_{m_1}\otimes^n)\otimes(V_{m_2}\otimes^{\lambda-n}) \tag{18}$$

The elements C_k S_λ which take care of the proper ordering of the factors in the tensorial products are the coset generators of the subgroup $S_n\otimes S_{\lambda-n}$ in S_λ Note

that this decomposition of $V_m \otimes^N$ is joined by a reduction of $[\gamma]^N \subset [GL(m_1+m_2)]^N$, namely,

$$
\begin{aligned}
[\gamma]^N V_m \otimes^N &= (\gamma V_m) \otimes^N \\
&= (\gamma_1 V_{m_1} \oplus \gamma_2 V_{m_2}) \otimes^N \\
&= \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \oplus C_k \{ (\gamma_1 V_{m_1}) \otimes^n \} \\
&\qquad \otimes \{ (\gamma_2 V_{m_2}) \otimes^{N-n} \} \\
&= \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \oplus C_k [\gamma_1]^n \otimes [\gamma_2]^{N-n} (V_{m_1} \otimes^n) \\
&\qquad \otimes (V_{m_2} \otimes^{N-n}) \\
&= \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \oplus \{ C_k [\gamma_1]^n \otimes [\gamma_2]^{N-n} C_k^{-1} \} \\
&\qquad \times \{ C_k (V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n}) \} \qquad (19)
\end{aligned}
$$

Hence,

$$
[\gamma_1 \oplus \gamma_2]^N = \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \oplus C_k [\gamma_1]^n \otimes [\gamma_2]^{N-n} C_k^{-1}, \qquad (20)
$$

where the summand is the restriction of $[\gamma_1 \oplus \gamma_2]^N$ to $C_k(V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n})$

If we choose the basis elements in each of the spaces $C_k(V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n})$ such that they are obtained from an original basis in $(V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n})$ by the permutations C_k it follows that all elements $C_k[\gamma_1] \otimes [\gamma_2]^{N-n} C_k^{-1}$ for

$$
k = 1, \cdots, \binom{N}{n}
$$

have the same matrix

$$
[\mathbf{D}(\gamma_1)]^n \otimes [\mathbf{D}(\gamma_2)]^{N-n} \qquad (21)
$$

This is easily proved since C_k is a linear operator which does not act on the transformation coefficients of $[\mathbf{D}(\gamma_1)]^n \otimes [\mathbf{D}(\gamma_2)]^{N-n}$ We thus arrive at a reduction of the matrix representation of $GL(m)$ carried by the complete tensor space $V_m \otimes^N$ under $GL(m_1+m_2)$

We must now look for the reduction of the irreducible representations $\langle \lambda \rangle$ of $GL(m)$ under $GL(m_1+m_2)$ To this end we reduce the spaces $V_{m_1} \otimes^n$ and $V_{m_2} \otimes^{N-n}$ according to Formula (7)

$$
V_{m_1} \otimes^n = \sum_{(\mu)\ \text{partitions of } n} \sum_{i=1}^{f_{[\mu]}} \oplus R_i^{(\mu)},
$$

$$
V_{m_2} \otimes^{N-n} = \sum_{(\nu)\ \text{partitions of } N-n} \sum_{j=1}^{f_{[\nu]}} \oplus R_j^{(\nu)}, \qquad (22)
$$

where $R_i^{(\mu)}$ is irreducible under $[GL(m_1)]^n$ and $R_j^{(\nu)}$ is irreducible under $[GL(m_2)]^{N-n}$ Therefore the tensor product space $R_i^{(\mu)} \otimes R_j^{(\nu)}$ is irreducible under $[GL(m_1)]^n \otimes [GL(m_2)]^{N-n}$ and, consequently, under $GL(m_1+m_2)$ [2] Since we just derived [see Expression (21)] that all spaces $C_k(V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n})$ span the same matrix representation of $GL(m_1+m_2)$, it follows that all spaces $C_k(R_i^{(\mu)} \otimes R_j^{(\nu)})$ for

$$
k = 1, \cdots, \binom{N}{n}, \qquad i = 1, \cdots, f_{[\mu]}, \qquad j = 1, \cdots, f_{[\nu]}
$$

are irreducible under $GL(m_1+m_2)$ and span the same matrix representation $\mathbf{D}^{(\mu)} \otimes \mathbf{D}^{(\nu)}$

Let us imagine that we place the $n_{(\mu)} \times n_{(\nu)}$ basis vectors of $R_i^{(\mu)} \otimes R_j^{(\nu)}$ in rows Taking these rows for $i = 1, \cdots, f_{[\mu]}$, $j = 1, \cdots, f_{[\nu]}$, we obtain a scheme comparable to Fig 1 We construct identical schemes for the spaces $C_k(R_i^{(\mu)} \otimes R_j^{(\nu)})$ and place these directly under the first scheme, getting

$$
\binom{N}{n}
$$

blocks with each $f_{[\mu]} \times f_{[\nu]}$ rows As shown above, each row carries the same irreducible matrix representation $\mathbf{D}^{(\mu)} \otimes \mathbf{D}^{(\nu)}$ of $GI(m_1+m_2)$ Using Weyl's result [Eq (7)] again, we find that each column in the first block spans the irreducible representation $[\mu] \otimes [\nu]$ of $S_n \otimes S_{N-n}$ The columns of the next blocks are generated from the first block by the coset generators C_k of $S_n \otimes S_{N-n}$ in S_N We can prove that a full column spans the induced representation of S_N denoted by $[\mu] \otimes [\nu] \uparrow S_N$ Every one of the $n_{(\mu)} \times n_{(\nu)}$ different columns spans an identical induced representation, thus assuring that all these induced representations can be reduced by the same basis transformation But, if we perform the same basis transformation on all columns, we do not disturb the matrix representation of $GL(m_1+m_2)$ afforded by the rows

Let the reduction of every column be written as follows

$$
[\mu] \otimes [\nu] \uparrow S_N = \sum_{(\lambda)\ \text{partitions of } N} m_{\mu\nu\lambda} [\lambda], \qquad (23)
$$

and let us perform on the columns the basis transformation corresponding to this reduction In our scheme we then have blocks in which the columns span the irreducible representations $[\lambda]$ of S_N, the same block appearing $m_{\mu\nu\lambda}$ times The rows still span the representation $\mathbf{D}^{(\mu)} \otimes \mathbf{D}^{(\nu)}$ of $GI(m_1+m_2)$ This result is

expressed mathematically as

$$\sum_{k=1}^{\binom{N}{n}} \sum_{i=1}^{f_{[\mu]}} \sum_{j=1}^{f_{[\nu]}} \oplus C_k(R_i^{(\mu)} \otimes R_j^{(\nu)}) = \sum_{(\lambda)}^{\text{partitions of } N} \oplus m_{\mu\nu\lambda} \sum_{r=1}^{f_{[\lambda]}} R_r^{(\mu)\otimes(\nu)} \quad (24)$$

Note that the space $R_r^{(\mu)\otimes(\nu)}$ cannot be written as a tensor product, although it spans the representation $D^{(\mu)} \otimes D^{(\nu)}$

We now consider the complete reduction of the full space $V_m \otimes^N$ under $GL(m_1+m_2)$

$$V_m \otimes^N = \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \oplus C_k(V_{m_1} \otimes^n) \otimes (V_{m_2} \otimes^{N-n})$$

$$= \sum_{n=0}^{N} \sum_{k=1}^{\binom{N}{n}} \sum_{(\mu)}^{\text{part of } n} \sum_{(\nu)}^{\text{part of } N-n} \sum_{i=1}^{f_{[\mu]}} \sum_{j=1}^{f_{[\nu]}} \oplus C_k R_i^{(\mu)} \otimes R_j^{(\nu)} \quad (25)$$

Substituting Formula (24) yields

$$V_m \otimes^N = \sum_{n=0}^{N} \sum_{(\mu)}^{\text{part of } n} \sum_{(\nu)}^{\text{part of } N-n} \sum_{(\lambda)}^{\text{part of } N} m_{\mu\nu\lambda} \times \sum_{r=1}^{f_{[\lambda]}} \oplus R_r^{(\mu)\otimes(\nu)} \quad (26)$$

Comparing this to the original reduction of $V_m \otimes^N$ under $GL(m)$,

$$V_m \otimes^N = \sum_{(\lambda)}^{\text{part of } N} \sum_{r=1}^{f_{[\lambda]}} \oplus R_r^{(\lambda)}, \quad (27)$$

we find

$$R_r^{(\lambda)} = \sum_{n=0}^{N} \sum_{(\mu)}^{\text{part of } n} \sum_{(\nu)}^{\text{part of } N-n} \oplus m_{\mu\nu\lambda} R_r^{(\mu)\otimes(\nu)} \quad (28)$$

This reduction is accompanied by a decomposition of the matrix representation

$$D^{(\lambda)}(\gamma_1 \oplus \gamma_2) = \sum_{n=0}^{N} \sum_{(\mu)}^{\text{part of } n} \sum_{(\nu)}^{\text{part of } N-n} \oplus m_{\mu\nu\lambda} D^{(\mu)}(\gamma_1) \otimes D^{(\nu)}(\gamma_2) \quad (29)$$

With this result we have fulfilled our aim the decomposition of the irreducible representations $\langle\lambda\rangle$ of $GL(m)$ with respect to $GL(m_1+m_2)$ They are expressed in tensor products of irreducible representations $\langle\mu\rangle$ and $\langle\nu\rangle$ of $GL(m_1)$ and $GL(m_2)$, respectively, with multiplicity coefficients that are derived from an induction problem in S_N.[23] According to Frobenius' reciprocity theorem we can also obtain the coefficients from the subduction

$$[\lambda] \downarrow S_n \otimes S_{N-n} = \sum_{(\mu)}^{\text{part of } n} \sum_{(\nu)}^{\text{part of } N-n} \oplus m_{\lambda\mu\nu}[\mu] \otimes [\nu], \quad (30)$$

where $m_{\mu\nu\lambda} = m_{\lambda\mu\nu}$

Generally this induction subduction problem is not multiplicity free i e $m_{\mu\nu\lambda}$ can be larger than 1 If we restrict the problem to many electron systems however only representations are allowed with Young diagrams of two columns at most, and $m_{\mu\nu\lambda}$ can just become 0 or 1 A constructive method of calculating these multiplicity factors has been given by Littlewood[24] and by Robinson [18] Extensive tables can be found in Appendix B of Ref 25

For demonstrating the application of our result (29) we again turn to the example of the O_2 molecule Suppose that we first wish to separate the core orbitals and then to perform an orbital transformation (e g hybridization or orthogonalization) which mixes only the valence orbitals of each atom among each other For the separation of the two core orbitals we consider the reduction of the triplet representation $\langle 2^7\ 1^3\rangle$ of $GL(10)$ under $GI\,(2+8)$

$$\langle 2^7, 1^2\rangle = \langle 1^\wedge\rangle \otimes \langle 2^7, 1\rangle$$

triplet O_2 (16 electrons) = doublet core (1 el) $\times$ doublet valence (15 el)

dimension 990 = 2 $\times$ 8

$$\oplus \langle 2\rangle \otimes \langle 2^6\ 1^2\rangle \oplus \langle 1^2\rangle \otimes \langle 2^7\rangle$$

singlet c (2) triplet v (14) + 3 $\times$ 28

triplet c (2) singlet v (14) + 1 $\times$ 36

$$\oplus \langle 1^2\rangle \otimes \langle 2^6\ 1^2\rangle \oplus \langle 2\ 1^2\rangle \otimes \langle 2^6\ 1^2\rangle$$

triplet c (2) triplet v (14) + 1 $\times$ 28

doublet c (3) doublet v (13) + 2 $\times$ 168

$$\oplus \langle 2\ 1^2\rangle \otimes \langle 2^5, 1^3\rangle \oplus \langle 2^2\rangle \otimes \langle 2^5\ 1^2\rangle$$

doublet c (3) quartet v (13) + 2 $\times$ 56

singlet c (4) triplet v (12) + 1 $\times$ 378

We are only interested in those states with 4 electrons in the core forming a singlet so that we are left with 12

electrons in the 8 valence orbitals carrying the triplet representation $\langle 2^6, 1^2\rangle$ of $GL(8)$. Formula (14) shows that we find 378 such states. The second step of our procedure amounts to the reduction of the latter representation under $GL(4+4)$, which reads (indicating the atoms by A and B)

$$\langle 2^6, 1^2\rangle = \langle 2, 1^2\rangle \otimes \langle 2^4\rangle$$

valence triplet — triplet A^{2+} — singlet B^{2-}
(12 electrons) — (4 el.) — (8 el.)
dimension 378 = 15 × 1

$$\oplus \langle 2, 1^3\rangle \otimes \langle 2^3, 1\rangle \oplus \langle 2^2, 1\rangle \otimes \langle 2^2, 1\rangle$$

quartet A^+ — doublet B^- — doublet A^+ — doublet B^-
(5) — (7) — (5) — (7)
$+ \quad 4 \times 4 \quad + \quad 20 \times 4$

$$\oplus \langle 2^3\rangle \otimes \langle 2^2, 1^2\rangle \oplus \langle 2^2, 1^2\rangle \otimes \langle 2^2, 1^2\rangle$$

singlet A — triplet B — triplet A — triplet B
(6) — (6) — (6) — (6)
$+ \quad 10 \times 6 \quad + \quad 6 \times 6$

$$\oplus \langle 2^2, 1^2\rangle \otimes \langle 2^3\rangle \oplus \langle 2^3, 1\rangle \otimes \langle 2^2, 1\rangle$$

triplet A — singlet B — doublet A^- — doublet B^+
(6) — (6) — (7) — (5)
$+ \quad 6 \times 10 \quad + \quad 4 \times 20$

$$\oplus \langle 2^3, 1\rangle \otimes \langle 2, 1^3\rangle \oplus \langle 2^4\rangle \otimes \langle 2, 1^2\rangle$$

doublet A^- — quartet B^+ — singlet A^{2-} — triplet B^{2+}
(7) — (5) — (8) — (4)
$+ \quad 4 \times 4 \quad + \quad 1 \times 15$

One could imagine that the secular problem including all configurations mixed by $GL(8)$ is too large. In this case, we can, for instance, consider only the triplet ground state which is formed by covalent interaction between the atomic ground state triplets. We then find 36 states which transform as

$$\langle 2^2\rangle \otimes \langle 2^2, 1^2\rangle \otimes \langle 2^2, 1^2\rangle$$

core — atom A — atom B
singlet — triplet — triplet

It is important to note, as we remarked before, that the product notation for the representations does not imply that the wavefunctions are simple tensorial products, they should have the proper symmetry also with respect to permutations exchanging electrons between subsystems. Still, they are basis elements of tensorial product representations. We have developed a method[26] to construct such wavefunctions by means of reduced Wigner operators, which is closely related to the aggregate theory of Matsen and Klein.[4]

VI. DISCUSSION

Starting from Weyl's theory which describes the behavior of antisymmetric spin eigenfunctions under general orbital transformations, we have derived how this behavior simplifies for partitioned orbital mixings. This simplification can be considerable, as in the example of the preceding section, where a full valence bond calculation is restricted to include only covalent states. Both the general result and its specification for partitioned transformations have numerous applications in quantum mechanical methods applied to atoms and molecules.

An example for the use of the general result is given by a complete orthogonalization of a linearly independent basis set. This orthogonalization of the orbital set can significantly change the interaction energy between atoms or molecules calculated with a limited number of atomic or molecular orbital configurations. In practice, this effect was noticed by Magnasco and Musso[27] in their computation of the interaction between two H_2 molecules and by Vonsovsky and Karpenko[28] in discussing superexchange by Anderson's model. The results of both studies depend sensitively on whether the atomic orbitals are orthogonalized or not. Weyl's theory tells in this case which configurations should be included in order to obtain a result which is independent of orbital mixing and, in case one takes fewer configurations, which new ones are introduced and how the weights are changed by a given orbital transformation.

Our special result for partitioned orbital transformations has been used in two different subjects so far.

(1) In connection with Roothaan's open shell Hartree Fock method[16] we have proved, very compactly, that

an antisymmetric spin eigenfunction is invariant under mixing of the closed shell orbitals,

an antisymmetric spin eigenfunction transforms under mixing of the open shell orbitals as if the closed shells were not present.

Using the latter theorem one shows very easily that by mixing of the open shells also non degenerate states (with the same spin multiplicity but different spatial symmetry) can be mixed.

(2) we have derived the explicit relation between a pair correlated DODS method and the CI approach. Particularly the effect that various matrix elements in the DODS secular equations do not depend on some mixing parameters could be explained.

These results will be elaborated in a forthcoming publication [25]

ACKNOWLEDGMENT

This research was supported by the Netherlands Organization for the Advancement of Pure Research (Z W O) via the Netherlands Foundation for Chemical Research (S O N)

[1] M Kotani, A Amemiya, E Ishiguro, and T Kimura, *Table of Molecular Integrals* (Maruzen, Tokyo 1955)
[2] D F Johnston, Rept Progr Phys **23**, 66 (1960)
[3] H Weyl, *The Theory of Groups and Quantum Mechanics* (Methuen, London, 1931)
[4] F A Matsen and D J Klein, J Phys Chem **75**, 1860 (1971) and references therein
[5] W A Goddard III Phys Rev **157**, 81 (1967)
[6] F E Harris, Advan Quantum Chem **3**, 61 (1967)
[7] R C Morrison and G A Gallup, J Chem Phys **50**, 1214 (1969)
[8] A van der Avoird, "Perturbation Theory for Intermolecular Forces," thesis, Eindhoven, 1968
[9] A T Amos and J I Musher, Chem Phys Letters **3**, 721 (1969)
[10] D J Klein, Intern J Quantum Chem **4S**, 271 (1971)
[11] R Eisenschitz and F London, Z Physik **60**, 491 (1930)
[12] H Boerner, *Darstellungen von Gruppen* (Springer, Berlin, 1967), 2nd ed
[13] F D Murnaghan, *The Theory of Group Representations* (Johns Hopkins U P , Baltimore 1938)
[14] R Pauncz, *Alternant Molecular Orbital Method* (Saunders, Philadelphia, 1967)
[15] A C Hurley, J E Lennard Jones, and J A Pople, Proc Roy Soc (London) **A220**, 446 (1953)
[16] C C J Roothaan, Rev Mod Phys **32**, 179 (1960)
[17] This is a group theoretical generalization of the theorem that two commuting operators have a set of common eigen functions In his proof of this relation (7), Weyl used the property that the enveloping algebra of $[GL(m)]^N$ is the commuting algebra of the group algebra of S_N Since the enveloping algebra of $[GL(m)]^N$ is also the enveloping algebra of $[U(m)]^N$, one finds that the (integral) representations $\{\lambda\}$ of the general linear group $GL(m)$ stay irreducible under the restriction of $GL(m)$ to $U(m)$ the group of unitary transformations (see Refs 12 and 13)
[18] G de B Robinson *Representation Theory of the Symmetric Group* (Edinburgh U P 1961)
[19] A J Coleman Advan Quantum Chem **4**, 83 (1968)
[20] F A Matsen J Am Chem Soc **92**, 3525 (1970)
[21] B R Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw Hill New York 1963)
[22] The first part of this assertion is a consequence of the theorem that the tensor products of the irreducible representations of two different groups form the irreducible representations of their outer product group This theorem was proved by Weyl (Ref 3, p 164) for the compact group $U(m)$ but it is also valid for integral representations of $GL(m)$ which stay irreducible under $U(m)$ (see Ref 17) The second statement holds because $[GL(m_1)]^n \otimes [GL(m_2)]^{N-n}$ is isomorphic to $GL(m_1) \otimes GL(m_2)$, which in its turn is isomorphic to $GL(m_1+m_2)$
[23] This result is a generalization of Weyl s branching theorem (Ref 3, p 391)
[24] D E Littlewood *The Theory of Group Characters* (Oxford U P New York 1950) 2nd ed
[25] B G Wybourne *Symmetry Principles and Atomic Spectroscopy* (Wiley–Interscience New York, 1970)
[26] P E S Wormer and A van der Avoird (unpublished)
[27] V Magnasco and G F Musso, J Chem Phys **48**, 2657 (1968)
[28] S V Vonsovsky and B V Karpenko *Encyclopedia of Physics, Vol 18/1 Magnetism*, edited by H P J Wijn (Springer, Berlin, 1968), p 265

Transformation Properties of Many-Electron Wave Functions with Special Attention to the Relation Between Pair-Correlated DODS and Configuration Interaction

P E S. WORMER AND A VAN DER AVOIRD
Institute of Theoretical Chemistry, University of Nijmegen, Toernooiveld, Nijmegen, The Netherlands

Abstracts

A method is presented that leads to a simple derivation of the explicit relation between pair-correlated DODS schemes (e g , the alternant molecular orbital method and the extended valence bond method) and configuration interaction

This result is based on a reduction formula for the representations of the general linear group, $GL(m)$, carried by the N-electron function space Generally, this paper deals with the effect of "partitioned" orbital transformations on states with "local" permutation symmetry

On presente une methode qui fournit une dérivation simple de la relation explicite entre les procédés de type DODS-paires correlées (par ex la methode des orbitales moléculaires alternantes et la methode de la mesomerie generalisee) et l'interaction des configurations

Ce résultat est basé sur une formule de réduction pour les représentations du groupe linéaire général, $GL(m)$, portees par l'espace des fonctions de N électrons En général l'article traite l'effet des transformations d'orbitales "partitionnées" sur les états avec une symétrie de permutation "locale"

Eine Methode wird vorgelegt, die zu einer einfachen Herleitung der expliziten Beziehung zwischen paar-korrelierten DODS-Verfahren (z B der Methode mit alternierenden Molekülorbitalen und der verallgemeinerten Valenzbindungsmethode) und Konfigurationswechselwirkung fuhrt

Das Ergebnis grundet sich auf eine Reduktionsformel fur die Darstellungen der allgemeinen linearen Gruppe, $GL(m)$, die von dem Raum der N-Elektronenfunktionen getragen werden Im allgemeinen behandelt der Artikel die Wirkung von "partionierten" Orbitaltransformationen auf Zustande mit "lokaler" Permutationssymmetrie

1. Introduction

Starting from a molecular calculation by the Hartree–Fock self-consistent field procedure, various methods exist for the computation of the remaining correlation energy One of the traditional methods is the configuration interaction technique, which permits the Hartree–Fock single configuration wave

function (which is an antisymmetrized N-electron eigenfunction of the total spin operator S^2 in restricted Hartree–Fock) to be mixed with excited configuration wave functions (of the same spin multiplicity). Excited configurations are constructed from the original configuration by promoting electrons from occupied to virtual Hartree–Fock orbitals. By Hartree–Fock orbitals we do not necessarily understand the canonical Hartree–Fock orbitals; they may localized, for instance, and are even not required to be orthogonal.

Other methods for calculating correlation energy follow the different orbitals for different spins (DODS) scheme described by Lowdin [1]. N different orbitals are obtained, for instance, by linear mixing between occupied and virtual Hartree–Fock orbitals. The correct N-electron wave function is formed from a product of such orbitals and a spin function product by antisymmetrization and spin projection. The mixing coefficients of the orbitals are considered as variation parameters which must be energy-optimized. Special forms of this DODS method mix each occupied Hartree–Fock orbital with just one virtual orbital, which should already account for a substantial part of the pair-correlation energy between the electrons in the doubly occupied Hartree–Fock orbital. Examples of such methods, which we shall call pair-correlated DODS, are given by the alternant molecular orbital method [2] and by the extended valence bond method [3, 4].

Although it can be readily recognized that the mixing between virtual and occupied orbitals introduces excited configurations into the Hartree–Fock wave function, so that DODS methods are in principle equivalent to configuration interaction with somewhat restricted interaction coefficients, it is not so easy to derive these coefficients explicitly. In a recent paper [5] we have reviewed the group-theoretical treatment of general orbital transformations as given by Weyl [6], and we have presented a new result which permits us to simplify the relation with CI considerably for pair-correlated DODS-methods. This is worked out in the present paper.

Our result is not restricted to the special orbital transformations occurring in pair-correlated DODS, and we have also indicated the specific transformation properties of some currently used N-electron wave functions.

2. Transformation Properties of N-Electron Wave Functions

We discuss N-electron wave functions that are composed of spatial orbitals and spin functions. They are considered to satisfy the Pauli principle, and to be eigenfunctions of the total spin operator S^2 with multiplicity $2S + 1$. Given a certain basis set of m orbitals we can construct $n_{\langle\lambda\rangle}$ linearly independent N-electron wave functions for a given S and S_z where $n_{\langle\lambda\rangle}$ is determined by m, N and S as described in ref. [5]. A linear transformation of the orbitals (e.g., a DODS mixing) will mix these N-electron wave functions among each other.

The formal theory for general orbital transformations was treated by Weyl [6] using the duality between the group $GL(m)$ of all linear (nonsingular) orbital

transformations and the permutation group S_N. This duality rests on the commutability of simultaneous orbital transformations for all N-electrons and the electron permutations. It implies that N-electron basis functions for a certain irreducible representation $[\lambda]$ of S_N are also basis functions for an irreducible representation of $GL(m)$, denoted by $\langle\lambda\rangle$. Now, we know that spatial wave functions which span a certain irreducible representation $[\lambda]$ of S_N, must combine with spin functions spanning the associate representation $[\tilde{\lambda}]$ (because of the antisymmetry of the total wave function) and, moreover, that this is uniquely connected to the spin functions being eigenfunctions of S^2 with definite S.

Combining these results, we find that the antisymmetric eigenfunctions of S^2 with fixed S form a basis for an irreducible representation $\langle\lambda\rangle$ of $GL(m)$. The irreducibility of $\langle\lambda\rangle$ implies that a general orbital transformation $\mathbf{D}(\gamma)$ (a nonsingular $m \times m$ matrix) will mix *all* $n_{\langle\lambda\rangle}$ N-electron antisymmetric wave functions belonging to the same S (and S_z). The mixing coefficients are the matrix elements of $\mathbf{D}(\gamma)^{\langle\lambda\rangle}$, the irreducible matrix representation $\langle\lambda\rangle$ of $GL(m)$ (dimension $n_{\langle\lambda\rangle}$). They are Nth order homogeneous polynomials in the matrix elements of $\mathbf{D}(\gamma)$, which can be derived after the actual construction of the N-electron basis functions of $\langle\lambda\rangle$.

Because orbital transformations do not affect the spin part of the wave functions, the theory is most easily formulated in a spinfree manner. The irreducible representations $[\lambda]$ of S_N, and also $\langle\lambda\rangle$ of $GL(m)$, are denoted by partitions of N or Young diagrams. The two-valuedness of the electron spin permits only two-row diagrams $[\frac{1}{2}N + S, \frac{1}{2}N - S]$ for the electron spin functions, so that the spatial wave functions must be basis elements of a two-column representation $[\lambda] = [2^{1/2N-S}, 1^{2S}]$ of S_N [7]. Such wave functions can be constructed by acting with the Wigner operators:

$$W_{jk}^{[\lambda]} = \frac{f_{[\lambda]}}{N!} \sum_{P \in S_N} D(P^{-1})_{kj}^{[\lambda]} P \tag{1}$$

on an N-fold orbital product which corresponds to the desired electron-configuration ($f_{[\lambda]}$ is the dimension of the irreducible representation $[\lambda]$ of S_N). By the theory of ref. [5] it is easily proved that the N-electron eigenfunctions of S^2 with given S, taken as a basis for the irreducible representations $\langle\lambda\rangle = \langle 2^{1/2N-S}, 1^{2S}\rangle$ of $GL(m)$, generate a matrix representation $\mathbf{D}(\gamma)^{\langle\lambda\rangle}$ that is exactly identical to the one which is generated by the spinfree basis projected by $W_{jk}^{[\lambda]}$ for arbitrary, but fixed j. We denote the latter basis for the representation $\langle\lambda\rangle$ of $GL(m)$ by:

$$\Psi_i^{\langle\lambda\rangle}(1, \cdots, N) = W_{jk}^{[\lambda]} \Phi_I(1, \cdots, N); \qquad i = 1, \cdots, n_{\langle\lambda\rangle} \tag{2}$$

with $$\Phi_I(1, \cdots, N) = \varphi_{i_1}(1)\varphi_{i_2}(2) \cdots \varphi_{i_N}(N)$$

The index I in principle runs over all orbital occupations ($i_1 = 1, \cdots, m; i_2 = 1, \cdots, m; \cdots; i_N = 1, \cdots, m$). Different choices of index k with fixed I project wave

functions which are either linearly independent or identical (except for a scalar factor which might be zero). The number of linearly independent wave functions that can be projected equals $n_{\langle\lambda\rangle}$.

Once we have constructed the spinfree basis functions of $\langle\lambda\rangle$ we can proceed with the explicit derivation of the elements of the matrix $\mathbf{D}(\gamma)^{\langle\lambda\rangle}$ by inspecting the result of an arbitrary orbital transformation $\mathbf{D}(\gamma)$ on these basis functions. If we realize that each orbital product in principle transforms into m^N products we can understand that this derivation is certainly not trivial. The complexity of the problem rapidly increases with N and m. For instance, a transformation of four orbitals in a four electron system mixes twenty singlet configurations with coefficients that are fourth order polynomials; for the singlet states of eight electrons in eight orbitals we need already a 1764 dimensional transformation matrix of eighth order polynomials. If we wish to find the explicit relation between a DODS treatment and a CI approach the expressions for the elements of these transformation matrices must be derived. Therefore it is very useful that we can simplify the transformation properties considerably if we restrict ourselves, for instance, to pair-correlated DODS.

3. Partitioned Orbital Transformations

The orbital transformations occurring in pair-correlated DODS are given by a direct sum of 2×2 matrices which mix every occupied Hartree-Fock orbital with one virtual orbital. They form a special case of "partitioned" orbital transformations, for which we have derived a simplified transformation formula for the N-electron wave functions. Partitioned orbital transformations are defined as follows:

$$\mathbf{D}(\gamma) = \mathbf{D}(\gamma_1) \oplus \mathbf{D}(\gamma_2) \tag{3}$$

where $\mathbf{D}(\gamma_1)$ mixes only m_1 orbitals and $\mathbf{D}(\gamma_2)$ mixes only m_2 orbitals among each other ($m_1 + m_2 = m$).

Transformations of this type form a group $GL(m_1 + m_2)$ which is a subgroup of $GL(m)$. We derived in ref. [5] that the restriction to this subgroup reduces the irreducible representation $\langle\lambda\rangle$ of $GL(m)$ in the following way:

$$\mathbf{D}(\gamma)^{\langle\lambda\rangle} \downarrow GL(m_1 + m_2) = \sum_{n=0}^{N} \sum_{\substack{(\mu) \\ \text{partitions} \\ \text{of } n}} \sum_{\substack{(\nu) \\ \text{partitions} \\ \text{of } N-n}} \oplus\, m_{\mu\nu\lambda} \mathbf{D}(\gamma_1)^{\langle\mu\rangle} \otimes \mathbf{D}(\gamma_2)^{\langle\nu\rangle} \tag{4}$$

where $\langle\mu\rangle$ are those irreducible representations of $GL(m_1)$, the group of transformations $\mathbf{D}(\gamma_1)$, labelled by partitions of n and ν those irreducible representations of $GL(m_2)$ labelled by partitions of $(N - n)$. The outer direct products $\langle\mu\rangle \otimes \langle\nu\rangle$ are irreducible representations of the outer product group $GL(m_1) \otimes GL(m_2)$, which is isomorphic to $GL(m_1 + m_2)$. The multiplicity coefficients

$m_{\mu\nu\lambda}$ are obtained from an induction problem in the permutation group, namely the induction of the representations of the product group $S_n \otimes S_{N-n}$ to the full permutation group S_N.

$$[\mu] \otimes [\ ] \uparrow S_N = \sum_{\substack{(\lambda) \\ \text{partitions} \\ \text{of } N}} \oplus\, m_{\mu\nu\lambda}[\lambda] \tag{5}$$

This reduction of the irreducible representations $\langle\lambda\rangle$ of $GL(m)$ for partitioned orbital transformations implies that no longer all $n_{\langle\lambda\rangle}$ basis functions of $\mathbf{D}(\gamma)^{\langle\lambda\rangle}$ are mixed among each other, but only the $n_{\langle\mu\rangle} \times n_{\langle\nu\rangle}$ basis functions of the product representations $\mathbf{D}(\gamma_1)^{\langle\mu\rangle} \otimes \mathbf{D}(\gamma_2)^{\langle\nu\rangle}$. The last number usually is considerably smaller and, moreover, the form of the mixing coefficients for N-electron wave functions will be much simpler.

In order to obtain these mixing coefficients explicitly we have to construct the N-electron basis functions of $\mathbf{D}(\gamma_1)^{\langle\mu\rangle} \otimes \mathbf{D}(\gamma_2)^{\langle\nu\rangle}$. These functions must simultaneously be basis elements of a two-column representation $[\lambda]$ of S_N. Wave functions which satisfy the latter condition (which is the spinfree equivalent of the total wave function satisfying the Pauli principle and being an eigenfunction of S^2), we shall call Pauli kets. The duality between $GL(m)$ and S_N implies that Pauli kets must also be basis elements of a two-column representation $\mathbf{D}(\gamma)^{\langle\lambda\rangle}$ of $GL(m)$ (see ref. [5]). Our special Pauli kets must therefore simultaneously be basis elements for $\langle\lambda\rangle$ of $GL(m)$ and for $\langle\mu\rangle$ & $\langle\nu\rangle$ of $GL(m_1 + m_2)$. We say that they are "sequence-adapted" to the two-membered sequence $GL(m) \supset GL(m_1 + m_2)$.

Let us start the construction of these Pauli kets by taking the basis functions for the irreducible representations $\mathbf{D}(\gamma_1)^{\langle\mu\rangle}$ and $\mathbf{D}(\gamma_2)^{\langle\nu\rangle}$, which are also basis functions of the representations $[\mu]$ of S_n and $[\nu]$ of S_{N-n}, respectively. (Again, because of the duality between the general linear group and the permutation group.) Such basis functions can be obtained by projection with the Wigner operators of S_n and S_{N-n} (compare formula (2)):

$$\begin{aligned} \Psi_p^{\langle\mu\rangle}(1, \quad , n) &= W_{qr}^{[\mu]}\Phi_r(1, \quad , n), & p &= 1, \cdots, n_{\langle\mu\rangle} \\ \Psi_s^{\prime\langle\nu\rangle}(n+1, \quad , N) &= W_{tu}^{[\nu]}\Phi_u(n+1, \quad \cdot, N), & s &= 1, \cdot \quad , n_{\langle\nu\rangle} \end{aligned} \tag{6}$$

The (tensorial) products of these basis functions

$$\{\Psi_p^{\langle\mu\rangle} \otimes \Psi_s^{\prime\langle\nu\rangle},\ p = 1, \quad , n_{\langle\mu\rangle},\ s = 1, \quad , n_{\langle\nu\rangle}\}$$

form a basis for the representations $\mathbf{D}(\gamma_1)^{\langle\mu\rangle}$ $\mathbf{D}(\gamma_2)^{\langle\nu\rangle}$ of $GL(m_1 + m_2)$. They are not Pauli kets, however, although they are adapted to the permutation symmetry of the group S_n S_{N-n}, because they do not span an irreducible representation of the full permutation group S_N. Pauli kets can be generated by acting with the left coset generators $\{\mathcal{C}_l,\ l = 1, \quad , (\ ^N_n)\}$ of the subgroup $S_n \subset S_{N-n}$ in S_N on the product functions, which yields a basis for the induced representation

$[\mu] \otimes [\nu] \uparrow S_N$, and a subsequent basis transformation which corresponds to the reduction (5). This induction-reduction process with respect to S_N does not influence the transformation properties of the wave functions, so that they still form a basis for $\mathbf{D}(\gamma_1)^{\langle\mu\rangle} \otimes \mathbf{D}(\gamma_2)^{\langle\nu\rangle}$. In fact, this construction forms an essential part of the derivation of formula (4), as described in ref. [5].

One way to obtain the proper Pauli kets is by direct action with the Wigner operators $W_{jk}^{[\lambda]}$ of the full permutation group S_N on the product functions $\Psi_p^{\langle\mu\rangle} \otimes \Psi_s^{\langle\nu\rangle}$. Thus, we can be sure of the correct permutation symmetry. It is better, however, to take advantage of the permutation symmetry which is already present in the product functions by using the "reduced Wigner operators":*

$$\tilde{W}_{JK}^{[\lambda]} = \frac{f_{[\lambda]}}{f_{[\mu]} f_{[\nu]}} \binom{N}{n}^{-1} \sum_{l=1}^{\binom{N}{n}} D(C_l^{-1})_{KJ}^{[\lambda]} C_l \tag{7}$$

$$\text{with } J = (\mu', \nu', v, w) \text{ and } K = (\mu, \nu, q, t)$$

In this formula $\mathbf{D}^{[\lambda]}$ must be an irreducible matrix representation of S_N which is "sequence adapted" to the subgroup $S_n \otimes S_{N-n}$. A method for deriving such representations is described by Matsen and Klein (ref. [8], Appendix 3). Acting with the reduced operators, which only contain the coset generators C_l of $S_n \otimes S_{N-n}$ in S_N, on a function which is already symmetry adapted to $S_n \otimes S_{N-n}$ we can obtain the same result as by acting with the full Wigner operator. This is expressed by the following relation:

$$W_{(\mu',\nu',v,w)(\mu,\nu,r,u)}^{[\lambda]} = \sum_{q=1}^{f_{[\mu]}} \sum_{t=1}^{f_{[\nu]}} \tilde{W}_{(\mu',\nu',v,w)(\mu,\nu,q,t)}^{[\lambda]} W_{qr}^{[\mu]} W_{tu}^{[\nu]} \tag{8}$$

The form of a sequence-adapted representation, the derivation of expression (8) and the meaning of all indices is shown in the Appendix.

Using this property of the reduced Wigner operators, we can write the resulting Pauli kets as follows:

$$\Psi_{p,s}^{\langle\mu\rangle\otimes\langle\nu\rangle}(1, \cdots, N) = \sum_{q=1}^{f_{[\mu]}} \sum_{t=1}^{f_{[\nu]}} \tilde{W}_{(\mu',\nu',v,w),(\mu,\nu,q,t)}^{[\lambda]} \Psi_p^{\prime\langle\mu\rangle}(1, \cdots, n) \otimes \Psi_s^{\prime\langle\nu\rangle}(n+1, \cdots, N) \tag{9}$$

where $p = 1, \cdots, n_{\langle\mu\rangle}$, $s = 1, \cdots, n_{\langle\nu\rangle}$ and the functions on the right-hand side are given by (6). The reduced Wigner operator being a linear combination of left coset generators, Equation (9) is the algebraic analogue of the induction–reduction process described before.

* This problem of obtaining the correct overall permutation symmetry for wave functions which are already permutation symmetry adapted to certain subsystems forms the basis of what is called by Matsen and Klein "the aggregate theory" [8, 9].

It is easily proved [5], by using the commutation of the matrices representing $GL(m_1 + m_2)$ and the coset generators C_i, that expression (9) indeed produces basis elements for the irreducible representation $\langle\mu\rangle \otimes \langle\nu\rangle$ of $GL(m_1 + m_2)$. Moreover, it produces proper Pauli kets with the representation $[\lambda]$ of S_N and $\langle\lambda\rangle$ of $GL(m)$, so that we now have obtained basis vectors which are adapted to the sequence $GL(m) \supset GL(m_1 + m_2)$. Equation (4) then shows the way to obtain a full basis of $GL(m)$.

The irreducibility of the representations $\langle\mu\rangle \otimes \langle\nu\rangle$ with respect to the group of partitioned orbital transformations $GL(m_1 + m_2)$ implies that the basis functions which we have now constructed give maximum simplification of the transformation matrix for the N-electron wave functions. Moreover, their expression by means of the reduced Wigner operators appears to be particularly convenient for deriving the remaining matrix elements explicitly. This we will demonstrate by an example in the next section.

We conclude this section by summarizing the above considerations in a corollary, while at the same time extending the theory to an arbitrary number of subsystems. Assume that the spatial wave function $\Phi^{(1)}(1, \cdots, n_1)$ is an orbital product, or a linear combination of orbital products, constructed from the first m_1 orbitals of the m-dimensional orbital basis. (It forms an element of the n_1-fold tensor product space $V_{m_1} \otimes^{n_1}$). Analogously, $\Phi^{(2)}(n_1 + 1, \cdots, n_1 + n_2)$ is constructed from the second set of m_2 orbitals. We continue until $\Phi^{(k)}(N - n_k + 1, \cdots, N)$, which is built from the last set of m_k orbitals.

$$(m_1 + m_2 + \cdots + m_\kappa = m;\ n_1 + n_2 + \cdots + n_k = N)$$

Let the Wigner operator $W_{JK}^{[\lambda]}$ be constructed on the basis of an irreducible representation $[\lambda]$ of S_N that is sequence adapted to

$$S_N \supset S_{n_1} \otimes S_{n_2} \otimes \cdots \otimes S_{n_k}$$

The column index K of the Wigner operator corresponds with the irreducible representation $[\mu_1] \otimes [\mu_2] \otimes \cdots \otimes [\mu_k]$ of the subgroup (see Appendix). By an easy generalization of the contents of this section one can show that the Pauli ket

$$\Psi = W_{JK}^{[\lambda]} \Phi^{(1)}(1, \cdots, n_1) \otimes \Phi^{(2)}(n_1 + 1, \cdots, n_1 + n_2) \otimes \cdots \otimes \Phi^{(k)}(N - n_k + 1, \cdots, N) \tag{10}$$

is an element in the basis of the irreducible representation $\langle\mu_1\rangle \otimes \langle\mu_2\rangle \otimes \cdots \otimes \langle\mu_k\rangle$ of the group $GL(m_1 + m_2 + \cdots + m_k)$. If the functions $\Phi^{(i)}$ are already symmetry adapted to the permutation group $S_{n_i} (i = 1, \cdots, k)$, we can use the reduced Wigner operators in order to replace (10) by an expression which is the generalization of (9) to an arbitrary number of subsystems.

By constructing basis functions that are sequence adapted to the permutation group sequence $S_N \supset S_{n_1} \otimes \cdots \otimes S_{n_k}$—Matsen and Klein [8] call such functions "aggregate states"—we have obtained a basis that is also sequence adapted to the orbital transformation groups $GL(m) \supset GL(m_1 + m_2 + \cdots + m_k)$. In other words, by fixing the parentage of the Pauli kets with respect to the permutation groups $S_{n_1}, S_{n_2}, \cdots, S_{n_k}$, we have completely determined their transformation behaviour under partitioned orbital transformations belonging to the group $GL(m_1 + m_2 + \cdots + m_k)$.

4. Relation between Pair-Correlated DODS and CI

The matrix of a pair-correlated DODS transformation is a direct sum of 2×2 matrices (not necessarily orthogonal), which mix pairs of orbitals that for physical reasons are chosen to be coupled. Matrix elements are the mixing coefficients which serve, according to the DODS formalism on a finite basis, as variation parameters. Mathematically, this pair-correlated DODS matrix is an element of a matrix group isomorphic to $GL(2 + 2 + \cdots + 2)$. In order to project Pauli kets transforming to irreducible representations of this group, one needs, according to the above prescription, representations $[\lambda]$ of S_N adapted to $S_2 \otimes S_2 \otimes \cdots \otimes S_2$. This type of representation is known as Serber representation [10]. The corollary of the previous section states that the DODS wave function projected by a Wigner operator on basis of a Serber representation transforms according to an irreducible representation of $GL(2 + 2 + \cdots + 2)$. Knowing this, we can readily derive the explicit expansion of the DODS wave function in terms of configurations, without necessity to write out the projection of any wave function involved in the expansion.

Let us illustrate the characteristic features of the theory by a rather simple example: the DODS treatment of a four electron system. Due to the simplifications which the theory permits for pair-correlated DODS, this treatment can easily be extended to larger systems.

We start with a basis set of four orbitals $\{\varphi_1, \varphi_3, \varphi_2, \varphi_4\}$, e.g., two occupied and two virtual molecular orbitals from a SCF–LCAO calculation. (In the AMO method these MO's are the canonical ones, in extended valence bond they are localized). The DODS scheme allows these four orbitals to be mixed in order to form four different orbitals $\{\psi_1, \psi_3, \psi_2, \psi_4\}$, in each of which one electron is placed. This can be achieved by a 4×4 orbital transformation matrix $\mathbf{D}(\gamma)$. The theory for general orbital transformations then tells us that we mix twenty singlet configurations: the SCF ground state with two doubly occupied MO's: $\varphi_1^2\varphi_2^2$ and all excited singlet configurations which can be constructed within the given basis. The mixing coefficients of the four-electron configurations, which form a 20×20 matrix, are fourth order homogeneous polynomials of the elements of $\mathbf{D}(\gamma)$.

If we restrict ourselves to pair-correlated DODS the matrix $\mathbf{D}(\gamma)$ has the following

typical form:

$$\mathbf{D}(\gamma) = \begin{pmatrix} d_{11} & d_{13} & 0 & 0 \\ d_{31} & d_{33} & 0 & 0 \\ 0 & 0 & d_{22} & d_{24} \\ 0 & 0 & d_{42} & d_{44} \end{pmatrix} = \mathbf{D}(\gamma_1) \oplus \mathbf{D}(\gamma_2) \tag{11}$$

In this example the occupied orbitals φ_1 and φ_2 are mixed with the virtual orbitals φ_3 and φ_4, respectively. Partitioned orbital transformations of this type cause the following reduction of the twenty-dimensional singlet representation of $GL(4)$ into outer direct products of representations of $GL(2)$:

$$\begin{aligned} \langle 2^2, 0^2 \rangle = \langle 2^2 \rangle \otimes \langle 0^2 \rangle \oplus \langle 2, 1 \rangle \otimes \langle 1, 0 \rangle \oplus \\ \oplus \langle 2, 0 \rangle \otimes \langle 2, 0 \rangle \oplus \langle 1^2 \rangle \otimes \langle 1^2 \rangle \oplus \\ \oplus \langle 1, 0 \rangle \otimes \langle 2, 1 \rangle \oplus \langle 0^2 \rangle \otimes \langle 2^2 \rangle \end{aligned} \tag{12}$$

with dimensionality $20 = 1 \times 1 + 2 \times 2 + 3 \times 3 + 1 \times 1 + 2 \times 2 + 1 \times 1$. In order to indicate clearly to which group the representations belong we have placed zeros for the nonoccurring rows in the Young diagrams. Note that the representation $\mathbf{D}(\gamma)^{\langle 0^2 \rangle}$ is the identity representation of $GL(2)$ and that the representation $\mathbf{D}(\gamma)^{\langle 1,0 \rangle}$ is identical to $\mathbf{D}(\gamma)$ itself.

Because in the DODS scheme we assign four different orbitals to the four electrons, thus placing two electrons in each orbital pair, the DODS four-electron singlets can only span the representations $\langle 2, 0 \rangle \otimes \langle 2, 0 \rangle$ or $\langle 1^2 \rangle \otimes \langle 1^2 \rangle$. (The other representations occurring in (12) correspond physically with one or two orbitals being doubly occupied.) In fact, from the configuration $\psi_1\psi_3\psi_2\psi_4$ we can construct two linearly independent singlet wave functions, one of which spans the representation $\langle 1^2 \rangle \otimes \langle 1^2 \rangle$, whereas the other one is a basis element of $\langle 2, 0 \rangle \otimes \langle 2, 0 \rangle$. In order to derive the explicit relation between the DODS wave functions and a CI basis we construct these wave functions and show their transformation behaviour.

In the first step we use the Wigner operators of the group S_2 :

$$\begin{aligned} W^{[2,0]}\psi_1\psi_3 = \tfrac{1}{2}(\psi_1\psi_3 + \psi_3\psi_1) = d_{11}d_{13}W^{[2,0]}\varphi_1\varphi_1 + \\ + d_{31}d_{33}W^{[2,0]}\varphi_3\varphi_3 + (d_{11}d_{33} + d_{13}d_{31})W^{[2,0]}\varphi_1\varphi_3 \\ W^{[1^2]}\psi_1\psi_3 = \tfrac{1}{2}(\psi_1\psi_3 - \psi_3\psi_1) = (d_{11}d_{33} - d_{13}d_{31})W^{[1^2]}\varphi_1\varphi_3 \end{aligned} \tag{13}$$

We note that for the triplet state several terms cancel. One proceeds analogously for the other electron pair.

The next step is to find an irreducible matrix representation $[2^2, 0^2]$ of S_4

which is sequence adapted to $S_2 \otimes S_2$. Accidentally, in this case the Serber representation is identical with the Young-Yamanouchi representation [11, 12]. We do not need this Serber representation explicitly, we only need to know the parentage of its different columns. The operator $W_{11}^{[2^2,0^2]}$ has the parentage $[2, 0] \otimes [2, 0]$, whereas $W_{12}^{[2^2,0^2]}$ corresponds with $[1^2] \otimes [1^2]$. Therefore, we can write the DODS wave functions as follows:

$$
\begin{aligned}
\Psi^{\langle 2,0\rangle\otimes\langle 2,0\rangle} &= W_{11}^{[2^2,0^2]}\psi_1\psi_3\psi_2\psi_4 \\
&= \tilde{W}_{11}^{[2^2,0^2]}(W^{[2,0]}\psi_1\psi_3) \otimes (W^{[2,0]}\psi_2\psi_4) \\
\Psi^{\langle 1^2\rangle\otimes\langle 1^2\rangle} &= W_{12}^{[2^2,0^2]}\psi_1\psi_3\psi_2\psi_4 \\
&= \tilde{W}_{12}^{[2^2,0^2]}(W^{[1^2]}\psi_1\psi_3) \otimes (W^{[1^2]}\psi_2\psi_4)
\end{aligned} \tag{14}
$$

Substituting the formula (13) and using relation (8) again, we can easily express the DODS wave functions in terms of the MO configurations that form the basis of a CI treatment (they should be Pauli kets also):

$$
\begin{aligned}
\Psi^{\langle 2,0\rangle\otimes\langle 2,0\rangle} = {} & d_{11}d_{13}d_{22}d_{24}W_{11}^{[2^2,0^2]}\varphi_1\varphi_1\varphi_2\varphi_2 \\
& + d_{11}d_{13}d_{42}d_{44}W_{11}^{[2^2,0^2]}\varphi_1\varphi_1\varphi_4\varphi_4 \\
& + d_{31}d_{33}d_{22}d_{24}W_{11}^{[2^2,0^2]}\varphi_3\varphi_3\varphi_2\varphi_2 \\
& + d_{31}d_{33}d_{42}d_{44}W_{11}^{[2^2,0^2]}\varphi_3\varphi_3\varphi_4\varphi_4 \\
& + d_{11}d_{13}(d_{22}d_{44} + d_{24}d_{42})W_{11}^{[2^2,0^2]}\varphi_1\varphi_1\varphi_2\varphi_4 \\
& + (d_{11}d_{33} + d_{13}d_{31}')d_{22}d_{24}W_{11}^{[2^2,0^2]}\varphi_1\varphi_3\varphi_2\varphi_2 \\
& + d_{31}d_{33}(d_{22}d_{44} + d_{24}d_{42})W_{11}^{[2^2,0^2]}\varphi_3\varphi_3\varphi_2\varphi_4 \\
& + (d_{11}d_{33} + d_{13}d_{31})d_{42}d_{44}W_{11}^{[2^2,0^2]}\varphi_1\varphi_3\varphi_4\varphi_4 \\
& + (d_{11}d_{33} + d_{13}d_{31})(d_{22}d_{44} + d_{24}d_{42})W_{11}^{[2^2,0^2]}\varphi_1\varphi_3\varphi_2\varphi_4 \\
\Psi^{\langle 1^2\rangle\otimes\langle 1^2\rangle} = {} & (d_{11}d_{33} - d_{13}d_{31})(d_{22}d_{44} - d_{24}d_{42})W_{12}^{[2^2,0^2]}\varphi_1\varphi_3\varphi_2\varphi_4
\end{aligned} \tag{15}
$$

This result tells us explicitly which configurations are mixed in a pair-correlated DODS wave function and how the mixing coefficients vary as a function of the orbital mixing coefficients.

In the special case of the AMO method, where the orbital transformation matrices have the form:

$$
\mathbf{D}(\gamma_1) = \begin{pmatrix} \cos\vartheta_1 & \cos\vartheta_1 \\ \sin\vartheta_1 & -\sin\vartheta_1 \end{pmatrix}, \qquad \mathbf{D}(\gamma_2) = \begin{pmatrix} \cos\vartheta_2 & \cos\vartheta_2 \\ \sin\vartheta_2 & -\sin\vartheta_2 \end{pmatrix} \tag{16}
$$

various terms cancel and we find the following simple relation:

$$\begin{aligned}\Psi^{\langle 2,0\rangle\otimes\langle 2,0\rangle} &= \cos^2\vartheta_1\cos^2\vartheta_2 W_{11}^{[2^2,0^2]}\varphi_1\varphi_1\varphi_2\varphi_2\\ &\quad-\cos^2\vartheta_1\sin^2\vartheta_2 W_{11}^{[2^2,0^2]}\varphi_1\varphi_1\varphi_4\varphi_4\\ &\quad-\sin^2\vartheta_1\cos^2\vartheta_2 W_{11}^{[2^2,0^2]}\varphi_3\varphi_3\varphi_2\varphi_2\\ &\quad+\sin^2\vartheta_1\sin^2\vartheta_2 W_{11}^{[2^2,0^2]}\varphi_3\varphi_3\varphi_4\varphi_4\\ \Psi^{\langle 1^2\rangle\otimes\langle 1^2\rangle} &= 4\cos\vartheta_1\sin\vartheta_1\cos\vartheta_2\sin\vartheta_2 W_{12}^{[2^2,0^2]}\varphi_1\varphi_3\varphi_2\varphi_4\end{aligned}$$

No singly or triply excited states occur in the singlet AMO wave functions.

This example can be easily extended to systems with more than four electrons by using the results of the previous section, formula (10) in particular. The complete pair-correlated DODS wave functions can be built up from the singlets and triplets of formula (13) by using the reduced Wigner operators based on the Serber representations of S_N. If one wants to derive which configurations occur in a certain DODS wave function and to calculate the configuration interaction coefficients as functions of the orbital mixing coefficients, it is not necessary to construct these complete wave functions, however. Only knowing their parentage with respect to $S_2 \otimes S_2 \otimes \cdots \otimes S_2$, i.e., the singlet and triplet states from which they are built, is sufficient.

5. Transformation of Some Common Antisymmetric Spin Eigenfunctions

We have seen in the previous section that the transformation properties of antisymmetric spin eigenfunctions under pair-correlated DODS mixing become particularly simple if we project such wave functions by means of Wigner operators based on the Serber representations of S_N. Although Serber functions have found recent interest in quantum chemistry [13], many other projectors have been used in the literature to construct antisymmetric eigenfunctions of S^2. We shall briefly discuss the transformation properties of some of these other Pauli kets also.

The general section about partitioned orbital transformations tells us that the transformation behaviour of Pauli kets can be largely simplified if one knows the permutational genealogy of their projectors.

Let us first look at some Wigner operators based on the orthogonal Young–Yamanouchi representations [12], called orthogonal units by Rutherford [11]. We shall label the basis functions for a given irreducible representation $[\lambda] = [2^{1/2N-S}, 1^{2S}]$ of S_N according to decreasing Yamanouchi symbols, using Hamermesh's [14] definition of these symbols. It follows immediately from the construction rules for the Young–Yamanouchi representations (ref. [14], Section 7-7) and from the form of the first standard Young tableau that the Wigner operator $W_{11}^{[\lambda]}$ is invariant under all transpositions (12), (34) $\cdots$, $(N-2S-1, N-2S)$

and spans an antisymmetric representation of all permutations among the last $2S$ electrons. Thus, the Wigner operator $W_{11}^{[\lambda]}$ generates spatial functions that carry the representation

$$\underbrace{[2] \otimes [2] \otimes \cdots \otimes [2]}_{\frac{1}{2}N - S \text{ factors}} \otimes [1^{2S}]$$

of the group $S_2 \otimes S_2 \otimes \cdots \otimes S_2 \otimes S_{2S}$. Such functions are also sequence adapted to

$$GL(m) \supset GL(m_1 + m_2 + \cdots m_{\frac{1}{2}N-S} + m_{\frac{1}{2}N-S+1})$$

and transform according to the representation:

$$\langle 2 \rangle \otimes \langle 2 \rangle \otimes \cdots \otimes \langle 2 \rangle \otimes \langle 1^{2S} \rangle \text{ of the latter group.}$$

In order to get nontrivial results the partitioning of the orbital basis ($m = m_1 + m_2 + \cdots + m_{1/2N-S+1}$) must satisfy the relations

$$m_1, m_2 \cdots, m_{1/2N-S} \geqq 1; \; m_{1/2N-S+1} \geqq 2S$$

This projector is used by Goddard in his GI-method [15].

Basis functions projected by this operator transform identically to one of the Serber basis functions under $GL(m_1 + m_2 + \cdots + m_{1/2N})$, namely according to

$$\underbrace{\langle 2 \rangle \otimes \cdots \otimes \langle 2 \rangle}_{\frac{1}{2}N - S \text{ factors}} \otimes \underbrace{\langle 1^2 \rangle \otimes \cdots \otimes \langle 1^2 \rangle}_{S \text{ factors}}$$

Consequently, this projector connects pair-correlated DODS *and* CI *in the manner* described in the previous section.

Another Wigner operator with a physically interesting parentage is $W_{f_{[\lambda]} \cdot f_{[\lambda]}}^{[\lambda]}$ where $f_{[\lambda]}$ is the dimension of the irreducible representation $[\lambda] = [2^{1/2N-S}, 1^{2S}]$. It is sequence-adapted to $S_N \supset S_{1/2N+S} \otimes S_{1/2N-S}$ and corresponds to the representation $[1^{1/2N+S}] \otimes [1^{1/2N-S}]$. Pauli kets which are projected by this operator transform according to $\langle 1^{1/2N+S} \rangle \otimes \langle 1^{1/2N-S} \rangle$ of $GL(m_1 + m_2)$. The partitioning of the orbital basis must satisfy $m_1 \geqq \frac{1}{2}N + S$, $m_2 \geqq \frac{1}{2}N - S$. If the equal signs hold, the irreducible representation $\langle 1^{1/2N+S} \rangle \otimes \langle 1^{1/2N-S} \rangle$ is one-dimensional, and the functions carrying $\langle 1^{1/2N+S} \rangle$ and $\langle 1^{1/2N-S} \rangle$ are simply Slater determinants composed of $\frac{1}{2}N + S$ and $\frac{1}{2}N - S$ spatial orbitals, respectively. In general, this projector corresponds to a physical system with antiferromagnetic coupling between two subsystems with $\frac{1}{2}N + S$ and $\frac{1}{2}N - S$ parallel spins. This projector is also of importance if one fills up degenerate DODS orbitals according to Hund's rules [16].

Also a common way of constructing antisymmetric eigenfunctions of S^2 is by acting with the Young operators [17], P_iN_i or N_iP_i, either on the spin part or on the spatial part of a wave function. .....

Each Young operator corresponds with a standard Young tableau ($i = 1, \cdots, f_{[\lambda]}$); the operator N_i is a product of antisymmetrizers over the electron indices occurring in the columns of tableau i, P_i symmetrizes over indices occurring in rows. These Young operators, which project nonorthogonal bases, can be related to the spinfree Wigner operators. In case they project spin functions, the first step involves transferring the permutational symmetry from spin to spatial coordinates by using one of the relations:

$$\begin{aligned} A(I \otimes P_iN_i^{[\lambda]}) &= A(P_{\tilde{i}}N_{\tilde{i}}^{[\tilde{\lambda}]} \otimes I) \\ A(I \otimes N_iP_i^{[\lambda]}) &= A(N_{\tilde{i}}P_{\tilde{i}}^{[\tilde{\lambda}]} \otimes I) \end{aligned} \tag{18}$$

where the first factor acts on the spatial coordinates and the second on the spin coordinates, I is the identity operator and A is the antisymmetrizer over spatial and spin coordinates. The representation $[\tilde{\lambda}]$ is associate to $[\lambda]$ and $\tilde{\imath} = f_{[\lambda]} + 1 - i$ counts according to increasing Yamanouchi symbols. These two relations are easily proved following the arguments of Heldmann [18]. Furthermore, one can show, using elements of Goddard's proof [19], that the following expansions must hold:

$$\begin{aligned} N_1P_1^{[\lambda]} &= \sum_{i=1}^{f_{[\lambda]}} c_iW_{i,1}^{[\lambda]} \\ P_{f_{[\lambda]}}N_{f_{[\lambda]}}^{[\lambda]} &= \sum_{i=1}^{f_{[\lambda]}} c_i'W_{i,f_{[\lambda]}}^{[\lambda]} \end{aligned} \tag{19}$$

Because the second index of the Wigner operators determines their permutational genealogy, and therewith the transformation properties of projected wave functions, and we have already discussed the Wigner operators $W_{11}^{[\lambda]}$ and $W_{f_{[\lambda]},f_{[\lambda]}}^{[\lambda]}$, relations (18) and (19) are sufficient to uncover these properties for the given Young operators.

The projector $N_{f_{[\lambda]}}P_{f_{[\lambda]}}^{[\lambda]}$ acting on a spin function $\alpha\beta\alpha\beta\cdots\alpha\beta\alpha\alpha\cdots\alpha$ generates the Boys–Reeves "spin-bonded" functions [20]. Using (18) and (19) we conclude that such functions transform according to $\langle 2\rangle \otimes \langle 2\rangle \otimes \cdots \otimes \langle 2\rangle \otimes \langle 1^{2S}\rangle$ under $GL(m_1 + m_2 + \cdots + m_{1/2N-S} + m_{1/2N-S+1})$. The projector $P_1N_1^{[\lambda]}$ acting on:

$$\underbrace{\alpha\alpha\cdots\alpha}_{\frac{1}{2}N+S}\underbrace{\beta\beta\cdots\beta}_{\frac{1}{2}N-S}$$

yields the same result as the Löwdin operator [21] acting on this function [22]. The Pauli kets corresponding to this choice transform as $\langle 1^{1/2N+S}\rangle \otimes \langle 1^{1/2N-S}\rangle$

under $GL(m_1 + m_2)$. If we have just N orbitals, these Pauli kets are invariant (except for a scalar factor) under mixing the spin-up and the spin-down orbitals among themselves. This result lies at the basis of the pairing theorem [23, 24].

6. Conclusion

In this paper we have treated the effect of orbital transformations on some commonly used many-electron wave functions, that are antisymmetric eigenfunctions of S^2. Especially, orbital mixings occurring in pair-correlated DODS-schemes were considered in more detail, in order to find the explicit relation between such schemes and the configuration interaction method. Orbital transformations were regarded as elements of the general linear group, $GL(m)$, and their effects were studied group theoretically. Although some of the transformation properties that we have discussed in this paper have already been derived for special cases [2, 25], the group theoretical approach places these results in a general and more formal framework. It gives much deeper insight as it clearly shows the relation between the permutational characteristics of wave functions and their transformation behaviour. Moreover, by using a group-theoretical formula that we have derived for the subduction of the representations of $GL(m)$, we can very often simplify complicated transformation matrices to a large extent and, thus, save much labour.

Appendix

We have shown in this paper that the transformation properties of N-electron wave functions under partitioned orbital transformations can actually be simplified if we construct basis functions of the type (9) by means of the "reduced" Wigner operators which satisfy relation (8). We will now show the derivation of this relation (8) and explain the meaning of all occurring indices. (A similar result has been derived by Klein et al. [9], using the theory of semi-simple algebras, and by Jahn [26] and Gerrat [16], who both restrict their attention to one- and two-column representations only).

Start with an irreducible representation $[\lambda]$ of S_N which is "sequence adapted" to the subgroup $S_n \otimes S_{N-n}$.

By definition such a representation $[\lambda]$ is completely decomposed if it is restricted to the subgroup $S_n \otimes S_{N-n}$. Therefore we can write for an arbitrary element PQ of this subgroup ($P \in S_n$, $Q \in S_{N-n}$):

$$\mathbf{D}(PQ)^{[\lambda]} = \sum_{\substack{(\mu) \\ \text{part. of } n}} \sum_{\substack{(\nu) \\ \text{part. of } N-n}} \oplus\, m_{\lambda\mu\nu} \mathbf{D}(P)^{[\mu]} \otimes \mathbf{D}(Q)^{[\nu]} \tag{A1}$$

(According to Frobenius' theorem the multiplicity coefficients in this subduction

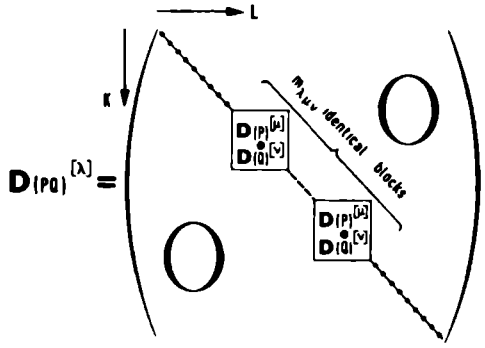


Figure 1. Form of a representation $[\lambda]$ which is adapted to the sequence $S_N \supset S_n \otimes S_{N-n}$.

problem are equal to those occurring in the induction (5):

$$m_{\lambda\mu\nu} = m_{\mu\nu\lambda})$$

The matrix $\mathbf{D}(PQ)^{[\lambda]}$ has a diagonal blocked form with the Kronecker products $\mathbf{D}(P)^{[\mu]} \otimes \mathbf{D}(Q)^{[\nu]}$ occurring $m_{\lambda\mu\nu}$ times on the diagonal (see Figure 1). If we want to write out relation (A1) in terms of matrix elements we necessarily have to introduce a large number of indices. The rows of the matrix $\mathbf{D}(PQ)^{[\lambda]}$ are numbered by:

μ (running over partitions of n)
ν (running over partitions of $N - n$),
$\alpha = 1, \cdots, m_{\lambda\mu\nu}$ (numbering multiple occurrences of $[\mu] \otimes [\nu]$),
$r = 1, \cdots, f_{[\mu]}$ (running over the rows of $\mathbf{D}(P)^{[\mu]}$),
$u = 1, \cdots, f_{[\nu]}$ (running over the rows of $\mathbf{D}(Q)^{[\nu]}$).

All these indices can be collected in one index K, numbering the rows of $\mathbf{D}(PQ)^{[\lambda]}$, such that there is a unique relation $K = (\mu, \nu, \alpha, r, u)$. In the same manner we label the columns by $L = (\mu', \nu', \alpha', q, t)$. Then, the following relation is obtained:

(A2)
$$D(PQ)^{[\lambda]}_{KL} = \delta_{\mu\mu'}\,\delta_{\nu\nu'}\,\delta_{\alpha\alpha'} D(P)^{[\mu]}_{rq} D(Q)^{[\nu]}_{ut}$$
$$\text{with } K = (\mu, \nu, \alpha, r, u) \text{ and}$$
$$L = (\mu', \nu', \alpha', q, t)$$

Although this factorization is only valid for the matrices $\mathbf{D}(PQ)^{[\lambda]}$ representing the subgroup $S_n \otimes S_{N-n}$, we must realize that the sequence-adapted representation $\mathbf{D}(R)^{[\lambda]}$ is defined for all $R \in S_N$. So, the labelling $K = (\mu, \nu, \alpha, r, u)$, $L = (\mu', \nu', \alpha', q, t)$ can be carried through for all $R \in S_N$.

We now write a Wigner operator for the (sequence-adapted) representation $[\lambda]$ of S_N :

$$W^{[\lambda]}_{JK} = \frac{f_{[\lambda]}}{N!} \sum_{R \in S_N} D(R^{-1})^{[\lambda]}_{KJ} R \tag{A3}$$

Every permutation $R \in S_N$ can be written as $R = C_l PQ$, where $P \in S_n$, $Q \in S_{N-n}$ and $\{C_l\ ;\ l = 1, \cdots, \binom{N}{n}\}$ are the (left) coset generators of $S_n \otimes S_{N-n}$ in S_N. Replacing the summation over all $R \in S_N$ by a threefold sum over $P \in S_n$, $Q \in S_{N-n}$ and $l = 1, \cdots, \binom{N}{n}$ we obtain

$$\begin{aligned} W^{[\lambda]}_{JK} &= \frac{f_{[\lambda]}}{N!} \sum_{P \in S_n} \sum_{Q \in S_{N-n}} \sum_{l=1}^{\binom{N}{n}} D(Q^{-1}P^{-1}C_l^{-1})^{[\lambda]}_{KJ} C_l PQ \\ &= \frac{f_{[\lambda]}}{N!} \sum_{P \in S_n} \sum_{Q \in S_{N-n}} \sum_{l=1}^{\binom{N}{n}} \sum_{L=1}^{f_{[\lambda]}} D(Q^{-1}P^{-1})^{[\lambda]}_{KL} D(C_l^{-1})^{[\lambda]}_{LJ} C_l PQ \end{aligned} \tag{A4}$$

Substitution of expression (7) for the reduced Wigner operator yields

$$W^{[\lambda]}_{JK} = \frac{f_{[\mu]} f_{[\nu]}}{n!\,(N-n)!} \sum_{L=1}^{f_{[\lambda]}} \tilde{W}^{[\lambda]}_{JL} \sum_{P \in S_n} \sum_{Q \in S_{N-n}} D(Q^{-1}P^{-1})^{[\lambda]}_{KL} PQ \tag{A5}$$

Since the permutation $Q^{-1}P^{-1}$ is an element of the group $S_n \otimes S_{N-n}$ we can use the decomposition formula (A2):

$$W^{[\lambda]}_{JK} = \frac{f_{[\mu]} f_{[\nu]}}{n!\,(N-n)!} \sum_{L=1}^{f_{[\lambda]}} \tilde{W}^{[\lambda]}_{JL} \sum_{P \in S_n} \sum_{Q \in S_{N-n}} \delta_{\mu\mu'} \delta_{\nu\nu'} \delta_{\alpha\alpha'} D(P^{-1})^{[\mu]}_{rq} D(Q^{-1})^{[\nu]}_{ut} PQ \tag{A6}$$

Remember that $K = (\mu, \nu, \alpha, r, u)$ and $L = (\mu', \nu', \alpha', q, t)$. Because of the Kronecker delta's the summation over $L = (\mu', \nu', \alpha', q, t)$, which is in fact a fivefold summation, can be reduced to a twofold sum. Moreover, we can substitute formula (1) for the Wigner operators of S_n and S_{N-n} to obtain the final result:

$$W^{[\lambda]}_{JK} = \sum_{q=1}^{f_{[\mu]}} \sum_{t=1}^{f_{[\nu]}} W^{[\lambda]}_{JL} W^{[\mu]}_{qr} W^{[\nu]}_{tu} \tag{A7}$$

where $K = (\mu, \nu, \alpha, r, u)$, $L = (\mu, \nu, \alpha, q, t)$ and $J = 1, \cdots, f_{[\lambda]}$ is some arbitrary index, which could also be written as a contraction of the same type: $J = (\mu', \nu', \alpha', v, w)$.

If we restrict ourselves to two-column representations of the permutation group, the induction-subduction problem (Equations (5) and (A1)) is multiplicity-free: $m_{\lambda\mu\nu}$ must be 0 or 1. In that case we can drop the indices α and α' to obtain Equation (8).

Formula (A7) can be easily generalized to an arbitrary number of subsystems. Suppose that we have n_1 electrons occupying the first m_1 orbitals, n_2 electrons in the second set of m_2 orbitals, etc., and n_k electrons in the last m_k orbitals, as in Equation (10). Irreducible representations of $S_{n_1}, S_{n_2}, \cdots, S_{n_k}$ are denoted by $[\mu_1], [\mu_2], \cdots, [\mu_k]$ and their rows and columns are numbered by $r_i = 1, \cdots, f_{[\mu_i]}$ and $q_i = 1, \cdots, f_{[\mu_i]}$, respectively, for $i = 1, \cdots, k$. The reduced Wigner operator $\tilde{W}^{[\lambda]}_{LK}$, which contains the left coset generators of the subgroup $S_{n_1} \otimes S_{n_2} \otimes \cdots \otimes S_{n_k}$ in S_N, is constructed on a representation $[\lambda]$ that is sequence-adapted to this subgroup. The rows and columns of $[\lambda]$ are numbered by $K = (\mu_1, \mu_2, \cdots, \mu_k, \alpha, r_1, r_2, \cdots, r_k)$ and $L = (\mu'_1, \mu'_2, \cdots, \mu'_k, \alpha', q_1, q_2, \cdots, q_k)$, respectively. The relation between a normal Wigner operator on the basis of this sequence-adapted representation $[\lambda]$ and the reduced Wigner operators becomes:

$$W^{[\lambda]}_{JK} = \sum_{q_1=1}^{f_{[\mu_1]}} \sum_{q_2=1}^{f_{[\mu_2]}} \cdots \sum_{q_k=1}^{f_{[\mu_k]}} \tilde{W}^{[\lambda]}_{JL} \tilde{W}^{[\mu_1]}_{q_1 r_1} W^{[\mu_2]}_{q_2 r_2} \cdots W^{[\mu_k]}_{q_k r_k} \tag{A8}$$

where $K = (\mu_1, \cdots, \mu_k, \alpha, r_1, \cdots, r_k)$, $L = (\mu_1, \cdots, \mu_k, \alpha, q_1, \cdots, q_k)$ and $J = (\mu'_1, \cdots, \mu'_k, \alpha', v_1, \cdots, v_k)$. This equation can be substituted into expression (10) in order to obtain a generalization of (9) for an arbitrary number of sub-systems.

Acknowledgment

Financial support from the Netherlands Organization for the Advancement of Pure Research (ZWO) via the Netherlands Foundation for Chemical Research (SON) is gratefully acknowledged.

The referee has drawn our attention to a recent paper [27] which is concerned with the transformation properties of N-electron wave functions under general orbital transformations (with rectangular matrices). Moreover, he mentioned that the equivalence between Young operators (PN or NP) and Lowdin projected determinants [19, 22] was also pointed out by Gallup [28].

Bibliography

[1] P. O. Lowdin, Phys. Rev. **97,** 1509 (1955).
[2] R. Pauncz, *The Alternant Molecular Orbital Method* (Saunders, Philadelphia, 1967).
[3] A. C. Hurley, J. E. Lennard-Jones, and J. A. Pople, Proc. Roy. Soc. (London) **A220,** 446 (1953).
[4] J. C. Slater, *Quantum Theory of Molecules and Solids*, Vol. I. (McGraw-Hill, New York, 1963).
[5] P. E. S. Wormer and A. van der Avoird, J. Chem. Phys. **57,** 2498 (1972).
[6] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Methuen, London, 1931; reprinted by Dover, New York, 1950).
[7] a. F. A. Matsen, J. Am. Chem. Soc. **92,** 2525 (1970);
b. J. I. Musher, J. Physique **31,** C4, 51 (1970).
[8] F. A. Matsen and D. J. Klein, J. Phys. Chem. **75,** 1860 (1971).

[9] D. J. Klein, C. H. Carlisle and F. A. Matsen, Advan. Quantum Chem. **5,** 219 (1970).
[10] R. Serber, Phys. Rev. **45.** 461 (1934).
[11] D. E. Rutherford, *Substitutional Analysis* (Edinburgh University Press, 1947).
[12] T. Yamanouchi, Proc. Phys.-Math Soc. Japan **19,** 436 (1937).
[13] W. I. Salmon and K. Ruedenberg, J. Chem. Phys. **57,** 2776 (1972).
[14] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[15] W. A. Goddard, Phys. Rev. **157,** 81 (1967).
[16] J. Gerratt in Atomic and Molecular Physics, Vol. 7, D. R. Bates and I. Estermann, ed., (Academic Press, New York, 1971), p. 141.
[17] D. J. Klein and B. R. Junker, J. Chem. Phys. **54,** 4290 (1971).
[18] G. Heldmann, Int. J. Quant. Chem. **2,** 785 (1968).
[19] W. A. Goddard, Phys. Rev. **169,** 120 (1968).
[20] C. M. Reeves, Comm. A.C.M. **9,** 276 (1966).
[21] P. O. Lowdin, Rev. Mod. Phys. **34,** 520 (1962).
[22] H. V. McIntosh, J. Math. Phys. **1,** 453 (1960).
[23] P. O. Lowdin, J. Appl. Phys. Suppl. **33,** 251 (1962).
[24] A. T. Amos and G. G. Hall, Proc. Roy. Soc. (London) **A263,** 483 (1961).
[25] M. Kotani, K. Ohno, and K. Kayama in Encyclopedia of Physics, Vol. 37/2, by S. Flugge, ed. (Springer, Berlin, 1961), p. 1.
[26] H. A. Jahn, Phys. Rev. **96,** 989 (1954).
[27] G. A. Gallup, Advan. Quantum Chem. **7,** 113 (1973), Section IIIB.
[28] G. A. Gallup, J. Chem. Phys. **50,** 1206 (1969).

Received September 4, 1973.
Revised December 25, 1973.

PART II:

INTERMOLECULAR FORCES

CHAPTER II 1 INTRODUCTION

The study of intermolecular forces impinges on many of the natural sciences. For example, intermolecular forces determine the thermodynamical properties of gases and liquids, they establish to a large extent the conformation and dynamical behaviour of the polymers, including the biopolymers, they control the structure of molecular crystals and biomembranes, etc. In all these examples one is mostly interested in the behaviour of the system at temperatures not much higher than room temperature, and consequently in the region of the intermolecular potential surface that reaches up to a few thousand calories per mole.

Calculations on such weak interactions are difficult mainly because of the following two reasons:

- The interactions are very small compared to the total energy of the system (the ratio is in the order of 1 to 10^6). One must realize in this respect that many quantum chemical methods compute interaction energies by subtracting energies that include the contributions of all the electrons in the system, including the inner shell electrons.
- The interaction itself is a delicate balance between a relatively strong repulsive and an equally strong attractive force.

The repulsive force has an exponential behaviour with an onset at a relatively short intermolecular distance R, therefore it is typically a short range force. The attractive force has a much longer range; accordingly the term long range force is frequently used as a synonym for intermolecular attraction. The physical origin of these forces has been explained qualitatively a long time ago, but it is only recently that quantitative calculations have come within reach.

The repulsive part of the potential,proposed by Van der Waals to account for the properties of imperfect gases,has been discussed by Eisenschitz and London [1]. They showed that the short range repulsion has the same quantum mechanical origin as chemical bonding, viz. the exchange phenomenon. The one-structure valence bond treatment of the hydrogen molecule is well-suited to illustrate this, because H_2 is a chemically bonded system in its singlet ground state, whereas the triplet state exhibits a repulsion which is characteristic for the complexes studied in this thesis. The H_2 singlet energy, then, is given by the formula [2]:

$$E_S(R) = \frac{C(R) + A(R)}{1 + S(R)}$$

and the triplet energy by:

$$E_T(R) = \frac{C(R) - A(R)}{1 - S(R)}.$$

The presence of the Coulomb term C(R) in these formulas can be rationalized quantum mechanically as well as classically. The exchange term A(R) (and also the overlap term S(R)) originates from the fact that electrons are indistinguishable, which means that the molecular wave function must be antisymmetric under permutation of all electrons. This typically quantum mechanical observation has been discussed in a more formal setting in sec. I.6.2 of this dissertation. Now, explicit calculation of A(R) and C(R) shows that A(R) is negative and much larger than C(R), at least for distances down to the equilibrium distance. Further $S(R) << 1$. So, A(R) gives rise to a strong (chemical) bonding in the singlet state, whereas the very same term causes a strong repulsion in the triplet state.

In principle it is feasible to exhibit the origin of the exchange repulsion also for larger systems than H_2, but because of the multitude of integrals arising in such an exhibition this would be very tedious and not very illuminating. However, all calculations performed to date bear out that the valence bond method (and also the Hartree-Fock method for that matter) always predicts exchange repulsion between systems without unsaturated valencies.

As a possible explanation for the long range intermolecular attraction several possibilities had already been proposed before the advent of quantum mechanics. Keesom [3] had attempted an explanation in terms of electrostatic interactions between permanent moments on the interacting systems. Debije [4] had suggested that induction might be an important factor. In 1930 London [5] came with a quantum mechanical treatment which unified into one single consistent theory the Keesom and Debije forces together with a force unknown until that time. London called this new force "dispersion force", because of the noticeable resemblance of his theory with the Kramers-Heisenberg dispersion theory. (It is interesting that a simple semi-classical version of London's theory [6], which considers an interacting dimer as two coupled harmonically oscillating dipoles, has also a strong likeness with the classical Drude-Lorentz theory of dispersion. Hence the name of the newly discovered force may be explained by the quantum mechanical as well as the classical analogy).

London's theory rests on two fundamental ideas: a Rayleigh-Schrödinger perturbation expansion of the interaction energy and a multipole expansion of the interaction operator. We will consider these points in somewhat more detail in the following chapters. This will be done because the calculations performed in the course of this work have been guided by London's theory, which therefore has considerable impact on this work. Furthermore, London's theory in its elegant irreducible tensor formulation, does not belong to the standard equipment of the average quantum chemist. Irreducible tensors allow, as we will see, the presentation of all the formulas in closed form.

The discussion of the long range intermolecular forces is followed by a few reprinted papers, the oldest of which presents computations on a charge transfer complex. Here the complete neglect of differential overlap (CNDO) method, a semi-empirical Hartree-Fock LCAO-MO method, has been employed. Although it is generally felt that the H-F model does not yield dispersion energy, it is accepted that the model is capable of giving electrostatic, induction and charge transfer interactions. So, reasonable answers may be obtained from H-F calculations on complexes in which charge transfer forces are thought to be the main bonding factor.

As for an approximation which neglects the differential overlap occurring in the integrals, it must be noted that in such a method the Coulomb integrals arising in the H-F equations do not vanish and so the CNDO method might in principle be able to give a reasonable account of the electrostatic forces. Moreover at the time the calculations were undertaken (autumn 1969) it was already known that CNDO usually gives fair results for hydrogen bonded complexes.

However, the paper shows rather unexpectedly that the transfer of charge from the donor to the acceptor is highly exaggerated, giving rise to unphysical interactions. Via a Mulliken population analysis the origin of these interactions could be traced back to the effect of charge transfer; at the same time this population analysis established also that no sort of intermolecular chemical bond was formed.

The other three articles reprinted below are chiefly based on the multi-structure valence bond method, although also CNDO and ab initio SCF calculations are reported for the ethylene dimer. The VB calculations were performed by computer programs especially developed for this purpose. (As a matter of fact the writing and debugging of these programs was the most time consuming part of the work performed for this dissertation). See

ref. 7 (reprinted in this thesis) for a short description of the programs. The basis in the VB calculations consists of canonical valence bond structures, defined in sec. I.5.4, and there called standard Weyl-Rumer tensors.

The unperturbed monomer states entering the VB calculations were all obtained from H-F calculations on the free monomers. To that end a standard SCF program (IBMOL-V) was employed. This program fulfilled all our needs, including the calculation of excited states. Because SCF-LCAO-MO calculations on open- and closed-shell systems have become a routine matter in most laboratories, and because the corresponding theory can be found in any text book on quantum chemistry, it is unnecessary to go into details about this part of the work. The same holds for the integration of atomic orbitals. See for instance ref. 8 for a discussion of these topics.

All the VB calculations have been performed within an orthogonalized basis of monomer orbitals. For He-He and the ethylene dimer we have investigated whether the intermolecular orthogonalization formed not too drastic an approximation. This was done by transforming the VB structures back to the original basis of non-orthogonal monomer orbitals. We used to that end matrices in irreps of GL(n). The explicit construction of such matrices has been described in sec. I.5.7 of this thesis.

CHAPTER II 2 THE MULTIPOLE EXPANSION OF $1/r_{12}$

In order to introduce the physical problem that forms the topic of this chapter, we consider a system of two point charges q_1 and q_2. We further take two coordinate systems: one with an origin at A and another one obtained by translation of the system at A along a vector $\vec{R}$. The second system has an origin at B. So, the vector $\vec{R}$ points from the origin of system A to the origin of system B. The coordinate vector of q_1 with respect to the system of axes at A is denoted by $\vec{r}_{1a}$; $\vec{r}_{1b}$ is the coordinate vector of q_1 with respect to the system of axes at B. So we have $\vec{r}_{1a} = \vec{r}_{1b} + \vec{R}$. The coordinate vectors of q_2 are analogously denoted by $\vec{r}_{2a}$ and $\vec{r}_{2b}$. Now, Coulomb's law states that the electrostatic interaction between q_1 and q_2 is proportional to:

$$1/r_{12} \equiv 1/|\vec{r}_{1a} - \vec{r}_{2a}| = 1/|\vec{r}_{1a} - (\vec{r}_{2b} + \vec{R})|.$$

If the length R of $\vec{R}$ is larger than $r_{1a} + r_{2b}$ the electrostatic interaction can be expanded into a series in 1/R. Each term in this series represents an interaction between a multipole centered on A and one centered on B; therefore this series is called the multipole expansion of the interaction. In this chapter this multipole expansion will be derived.

Obviously, the multipole expansion is an unnecessarily complicated way of expressing the interaction between two point charges. However, if we have a continuous charge distribution ρ_a around A and another continuous charge distribution ρ_b around B, then the multipole expansion furnishes a convenient means of computing the interaction between ρ_a and ρ_b. For in that case we let q_1 stand for $\rho_a(\vec{r}_{1a})dx_a dy_a dz_a$ and q_2 for $\rho_b(\vec{r}_{2b})dx_b dy_b dz_b$, expand the interaction and integrate term-wise over both the charge distributions. Provided R is always larger than $r_{1a} + r_{2b}$, that is, provided ρ_a and ρ_b do not overlap, we have found a converging and physically appealing manner of expanding the interaction between ρ_a and ρ_b.

Quantum mechanically the multipole expansion of $1/r_{12}$ enables the expansion of the interaction part of a Hamiltonian describing a system of two non-penetrating atoms or molecules.

There are two essentially different ways of writing the multipole series. The first employs Cartesian coordinate systems at the origins A and B and expresses the interaction completely in terms of Cartesian tensors. The "geometrical" variables do not appear in closed form but in

operator expressions of the type [9]:

$$T_{\alpha\beta\ldots\gamma} \equiv \nabla_\alpha \nabla_\beta \ldots \nabla_\gamma \left(\frac{1}{R}\right).$$

The operators ∇_α, ∇_β, ..., ∇_γ standing for the Cartesian components of the nabla operator, $T_{\alpha\beta\ldots\gamma}$ is a Cartesian tensor. The Cartesian approach has the disadvantage that the final formula is highly redundant. This follows from the fact that the components of Cartesian tensors span reps and not irreps of the full rotation group SO(3), and accordingly the problem has not been disentangled to its limit. (The group SO(3) arises in this connection as the invariance group of $1/r_{12}$). Rotational invariance is of course preserved upon expansion, and so the terms in the multipole expansion are all invariant under rotation. Reduction of the tensor components under SO(3) will indeed bring a considerable simplification, because out of the $\frac{1}{2}(n + 1)(n + 2)$ different components of a Cartesian multipole moment tensor of rank n only $2n + 1$ contribute to the interaction.

The second manner of writing the multipole expansion makes use of irreducible tensors, thus assuring a maximal disentanglement of the problem. An additional advantage of this approach is that the expansion is not only in closed form, but also solely in terms of spherical harmonics and Clebsch-Gordan coefficients, which can both be found in many tables. Furthermore we can often profit by the fact that the Wigner-Eckart theorem can be applied directly in ensuing calculations.

A disadvantage of the irreducible tensor notation is that the derivation of the multipole expansion, as for instance given by Rose [10], is rather involved, whereas the Cartesian derivation can be followed by everybody who knows how to differentiate a function of x, y and z. However, a considerably shorter derivation will be presented subsequently, which, as far as I am aware, has not been given before. The final result is presented in the form of a theorem (the second theorem of this chapter).

Firstly, we establish some notation and some necessary definitions. In the following we will adhere to Edmonds [11] with regard to all conventions, including for instance the phases of Clebsch-Gordan coefficients and of spherical harmonics, the definition of standard bases for irreps of SO(3), etc.

Definitions.

1. The function $Y_{\ell,m}(\hat{r})$ is a spherical harmonic of the order ℓ. See for an explicit definition ref. 11, formula 2.5.5. The vector $\hat{r}$ is a unit vector with the polar coordinates θ and ϕ.

2. The function:

$$S_{\ell,m}(\vec{r}) \equiv r^{\ell} Y_{\ell,m}(\hat{r})$$

is a regular solid harmonic.

The same function considered as a multiplicative operator is a multipole operator.

3. The function:

$$Z_{\ell,m}(\vec{r}) = r^{-\ell-1} Y_{\ell,m}(\hat{r})$$

is an irregular solid harmonic.

4. Let the set

$$\vec{T}_{\ell}(\vec{r}) \equiv \{T_{\ell,m}(\vec{r}) \,|\, m = -\ell, \ldots, +\ell\}$$

consist of $2\ell + 1$ quantities defined with respect to a set of coordinates in R^3. Describe a rotation of R^3 by:

$$\vec{r}' = \underline{R}\,\vec{r}$$

and let $\underline{D}^{(\ell)}(\underline{R})$ be the matrix representing $\underline{R}$ in the $(2\ell+1)$-dimensional irrep of SO(3). If the set $\vec{T}_{\ell}$ satisfies the transformation equations:

$$T_{\ell,m}(\vec{r}') = \sum_{m'=-\ell}^{+\ell} T_{\ell,m'}(\vec{r})\, D^{(\ell)}(\underline{R}^{-1})_{m'm}, \qquad m = -\ell, \ldots, +\ell$$

then $\vec{T}_{\ell}(\vec{r})$ is an irreducible tensorial set of the order ℓ. Examples of irreducible tensorial sets are spherical harmonics, angular momentum operators, multipole operators, etc.

5. Let $\vec{U}_{\ell_1}$ and $\vec{V}_{\ell_2}$ be irreducible tensorial sets. The irreducible product of these sets is defined by

$$[\vec{U}_{\ell_1} \times \vec{V}_{\ell_2}]^{(\ell)}_m \equiv \sum_{m_1,m_2} U_{\ell_1,m_1} V_{\ell_2,m_2} (\ell_1,m_1;\ell_2,m_2 \,|\, \ell,m), \qquad m = -\ell,\ldots,+\ell$$

where $(\ell_1,m_1;\ell_2,m_2\,|\,\ell m)$ is a Clebsch-Gordan coefficient. Clearly an irreducible product is an irreducible tensorial set.

6. Inserting the value of $(\ell,m;\ell,-m\,|\,0,0)$ one gets:

$$[\vec{U}_{\ell} \times \vec{V}_{\ell}]^{(0)}_0 = (-1)^{\ell}(2\ell+1)^{-\frac{1}{2}} \sum_m (-1)^m U_{\ell,-m} V_{\ell,m}.$$

7. Let $\vec{r} = (x,y,z)$ be a vector in R^3. Its spherical coordinates are defined by:

$$r_{+1} = -1/\sqrt{2}\,(x + iy) \qquad r_0 = z \qquad r_{-1} = 1/\sqrt{2}\,(x - iy)$$

In the sequel spherical coordinates of vectors will be labelled by m or μ.

8. The inner product of two vectors $\vec{r}$ and $\vec{r}'$ can be written thus:

$$\vec{r}\cdot\vec{r}' = \sum_m (-1)^m r_{-m} r'_m = -\sqrt{3}\,[\vec{r} \times \vec{r}']^{(0)}_0$$

The following lemma will be needed in the proof of the translation formula for irregular solid harmonics.

Lemma.

$$[\ldots[[\vec{r} \times \vec{r}]^{(2)} \times \vec{r}]^{(3)} \times \vec{r}]^{(4)} \ldots \times \vec{r}]^{(k)}_m = \sqrt{4\pi}\left[\frac{2^k}{(2k+1)!}\right]^{1/2} k!\, S_{k,m}(\vec{r}).$$

Proof. First note that $r_m = \sqrt{4\pi/3}\, S_{1,m}(\vec{r})$, m = -1, 0, +1, and therefore:

$$[\ldots[\vec{r} \times \vec{r}]^{(2)} \times \vec{r}]^{(3)} \ldots \times \vec{r}]^{(k)}_m = (\sqrt{4\pi/3})^k [\ldots[\vec{S}_1 \times \vec{S}_1]^{(2)} \times \vec{S}_1]^{(3)} \ldots \times \vec{S}_1]^{(k)}_m.$$

The right hand side can be easily evaluated by recursion from the relation:

$$[\vec{S}_j \times \vec{S}_1]^{(j+1)}_m = \sqrt{3/4\pi}\left[\frac{j+1}{(2j+3)}\right]^{1/2} S_{j+1,m}$$

which we now derive:

$$[\vec{S}_j \times \vec{S}_1]^{(j+1)}_m \equiv \sum_{m_1,m_2} (j,m_1;1,m_2 \,|\, j+1,m)\; S_{j,m_1} S_{1,m_2}.$$

Insert into the right hand side the Gaunt series (ref. 11, formula 5.1.6):

$$[\vec{S}_j \times \vec{S}_1]^{(j+1)}_m = \sum_{m_1,m_2} (j,m_1;1,m_2 \,|\, j+1,m) \sum_{\ell,m'} (j,m_1;1,m_2 \,|\, \ell,m') \times$$

$$\times\, (j,0;1,0\,|\,\ell,0) \times \left[\frac{3(2j+1)}{4\pi(2j+3)}\right]^{1/2} S_{\ell,m'}.$$

Using the fact that the C-G coefficients constitute a real orthogonal matrix and inserting

$$(j,0;1,0\,|\,j+1,0) = \left[\frac{j+1}{2j+1}\right]^{1/2}$$

we arrive at:

$$[\vec{S}_j \times \vec{S}_1]^{(j+1)}_m = \sqrt{3/4\pi}\left[\frac{j+1}{2j+3}\right]^{1/2} S_{j+1,m}.$$

If we apply this relation repeatedly, for j = 1, ..., k - 1, we get:

$$[\ldots[\vec{S}_1 \times \vec{S}_1]^{(2)} \times \vec{S}_1]^{(3)} \ldots \times \vec{S}_1]^{(k)}_m = (\sqrt{3/4\pi})^{k-1}\left[\frac{3\cdot k!\cdot k!\cdot 2^k}{(2k+1)!}\right]^{1/2} S_{k,m}.$$

Using $S_{1,m} = \sqrt{3/4\pi}\, r_m$ we obtain immediately the relation to be proved.

The now following theorem gives a translation formula for irregular solid harmonics.

Theorem.

$$Z_{\ell,m}(\vec{r}_1 + \vec{r}_2) = \sqrt{4\pi} \sum_{k=0}^{\infty} \binom{2\ell + 2k}{2k}^{\frac{1}{2}} (2k + 1)^{-\frac{1}{2}} [\vec{Z}_{\ell+k}(\vec{r}_2) \times \vec{S}_k(\vec{r}_1)]_m^{(\ell)}.$$

Proof. Expand according to Taylor:

$$Z_{\ell,m}(\vec{r}_1 + \vec{r}_2) = \sum_{k=0}^{\infty} \frac{1}{k!} (\vec{r}_1 \cdot \vec{\nabla}_2)^k Z_{\ell,m}(\vec{r}_2).$$

To evaluate the terms in the expansion we use the gradient formula (ref. 12, p. 150):

$$\sum_{\mu} (-1)^{\mu} r_{1,-\mu} \nabla_{2,\mu} Z_{\ell,m}(\vec{r}_2) = [(2\ell + 1)(\ell + 1)]^{\frac{1}{2}} \sum_{\mu} (\ell+1, m+\mu; 1, -\mu \,|\, \ell, m) \times \times Z_{\ell+1,m+\mu}(\vec{r}_2)\, r_{1,-\mu},$$

which written more compactly reads:

$$\vec{r}_1 \cdot \vec{\nabla}_2 \, Z_{\ell,m}(\vec{r}_2) = [(2\ell + 1)(\ell + 1)]^{\frac{1}{2}} [\vec{Z}_{\ell+1}(\vec{r}_2) \times \vec{r}_1]_m^{(\ell)}.$$

Repeated application of $\vec{r}_1 \cdot \vec{\nabla}_2$ gives:

$$(\vec{r}_1 \cdot \vec{\nabla}_2)^k Z_{\ell,m}(\vec{r}_2) = c_{k\ell m} [\ldots [\vec{Z}_{\ell+k}(\vec{r}_2) \times \vec{r}_1]^{(\ell+k-1)} \times \vec{r}_1]^{(\ell+k-2)} \times \ldots \vec{r}_1]_m^{(\ell)}$$

and one easily derives for the constant $c_{k\ell m}$:

$$c_{k\ell m} = \left[\frac{(2\ell + 2k)!}{(2\ell)!\, 2^k}\right]^{\frac{1}{2}}.$$

The coupling scheme originating from the application of $(\vec{r}_1 \cdot \vec{\nabla}_2)^k$ is equivalent to the quantum mechanical coupling case of adding k particles, all with quantum number $j = 1$, to a state with $j = \ell + k$ in such a manner that a state with $j = \ell$ results. This same state can also be obtained by first coupling the k particles to a state of maximal multiplicity, i.e. with $j = k$, and then coupling this state with $j = \ell + k$ to give $j = \ell$. This is so because there is only one state with quantum numbers $j = \ell$ and $j_z = m$ and genealogy $(\ell+k, 1, 1, \ldots, 1 \,|\, \ell)$. So:

$$[\ldots [\vec{Z}_{\ell+k}(\vec{r}_2) \times \vec{r}_1]^{(\ell+k-1)} \times \vec{r}_1] \ldots \times \vec{r}_1]_m^{(\ell)} = = [\vec{Z}_{\ell+k}(\vec{r}_2) \times [\ldots [\vec{r}_1 \times \vec{r}_1]^{(2)} \times \vec{r}_1]^{(3)} \ldots \times \vec{r}_1]^{(k)}]_m^{(\ell)}.$$

From the lemma just proved:

$$[\ldots[\vec{r}_1 \times \vec{r}_1]^{(2)} \times \vec{r}_1]^{(3)} \ldots \vec{r}_1]_m^{(k)} = \sqrt{4\pi}\left[\frac{2^k}{(2k+1)!}\right]^{\frac{1}{2}} k!\, S_{k,m}(\vec{r}_1)$$

and so:

$$(\vec{r}_1\cdot\vec{\nabla}_2)^k Z_{\ell,m}(\vec{r}_2) = \sqrt{4\pi}\, k! \binom{2\ell+2k}{2k}^{\frac{1}{2}} (2k+1)^{-\frac{1}{2}} [\vec{Z}_{\ell+k}(\vec{r}_2) \times \vec{S}_k(\vec{r}_1)]_m^{(\ell)}.$$

Substitution of this expression into the Taylor expansion gives the desired result.

We have now at our disposal the necessary machinery to state and prove the essential result of this chapter.

Theorem (The multipole expansion of $1/r_{12}$).

$$1/r_{12} = (4\pi)^{3/2} \sum_{\ell_a=0}^{\infty} \sum_{\ell_b=0}^{\infty} (-1)^{\ell_b} \binom{2L}{2\ell_a}^{\frac{1}{2}} [(2\ell_a+1)(2\ell_b+1)(2L+1)]^{-\frac{1}{2}} \times$$

$$\times \sum_{M=-L}^{+L} (-1)^M Z_{L,-M}(\vec{R})\, [\vec{S}_{\ell_b}(\vec{r}_{2b}) \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_M^{(L)}.$$

Here: $L = \ell_a + \ell_b$; $\vec{r}_{1a}$, $\vec{r}_{2b}$ and $\vec{R}$ are the vectors introduced in the beginning of this chapter, and the lengths of the vectors satisfy: $R > r_{1a} + r_{2b}$.

Proof. Apply the Laplace expansion of $1/r_{12}$ [13, p. 79]:

$$1/r_{12} = \sum_{\ell_a=0}^{\infty} 4\pi(2\ell_a+1)^{-\frac{1}{2}} (-1)^{\ell_a} [\vec{Z}_{\ell_a}(\vec{r}_{2a}) \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_0^{(0)}.$$

Because, as discussed in the beginning of this chapter, $R > r_{1a} + r_{2b}$ it follows that always $r_{2a} > r_{1a}$. This has been used in the Laplace expansion just given. Substituting $\vec{r}_{2a} = \vec{r}_{2b} + \vec{R}$, and applying the foregoing theorem we find:

$$1/r_{12} = (4\pi)^{3/2} \sum_{\ell_a=0}^{\infty} \sum_{\ell_b=0}^{\infty} (-1)^{\ell_a} [(2\ell_a+1)(2\ell_b+1)]^{-\frac{1}{2}} \binom{2\ell_a+2\ell_b}{2\ell_a}^{\frac{1}{2}} \times$$

$$\times [[\vec{Z}_{\ell_a+\ell_b}(\vec{R}) \times \vec{S}_{\ell_b}(\vec{r}_{2b})]^{(\ell_a)} \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_0^{(0)}.$$

There is only one state with $j = 0$, $j_z = m$ and genealogy $(\ell_a+\ell_b, \ell_b, \ell_a | 0)$ and so we can write, substituting $L \equiv \ell_a + \ell_b$:

$$[[\vec{Z}_L(\vec{R}) \times \vec{S}_{\ell_b}(\vec{r}_{2b})]^{(\ell_a)} \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_0^{(0)} = [\vec{Z}_L(\vec{R}) \times [\vec{S}_{\ell_b}(\vec{r}_{2b}) \times \vec{S}_{\ell_a}(\vec{r}_{1a})]^{(L)}]_0^{(0)} =$$

$$= (-1)^L (2L+1)^{\frac{1}{2}} \sum_M (-1)^M Z_{L,-M}(\vec{R}) [\vec{S}_{\ell_b}(\vec{r}_{2b}) \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_M^{(L)}.$$

(Of course one can prove the validity of this recoupling more explicitly by evaluation of the associated 6j-symbol).
Substituting this result in the expansion of $1/r_{12}$ we have proved the theorem.

Notes.

1. The expression $[\vec{S}_{\ell_b}(\vec{r}_{2b}) \times \vec{S}_{\ell_a}(\vec{r}_{1a})]_M^{(L)}$ represents physically the coupling of a multipole moment of order ℓ_b on B with a multipole moment of order ℓ_a on A to a total moment of maximal order L.
2. The terms in the expansion are invariant under SO(3) as it, of course, must be.
3. The multipole expansion in this form clearly separates the "geometrical" variables $\vec{R}$ from the "structure" variables contained in the irreducible product of moments.
4. To exhibit more clearly the 1/R-dependence of the terms, we use the fact that the range $0 \leq \ell_a < \infty$, $0 \leq \ell_b < \infty$ is also covered by $0 \leq L < \infty$, $0 \leq \ell_b \leq L$. Eliminating ℓ_a we get:

$$1/r_{12} = (4\pi)^{3/2} \sum_{L=0}^{\infty} R^{-L-1} Y_{L,-M}(\hat{R}) \times$$

$$\times \{ \sum_{\ell_b=0}^{L} \sum_{M=-L}^{+L} (-1)^M c_{L,\ell_b} [\vec{S}_{\ell_b}(\vec{r}_{2b}) \times \vec{S}_{L-\ell_b}(\vec{r}_{1a})]_M^{(L)} \}$$

where:

$$c_{L,\ell_b} = (-1)^{\ell_b} \begin{pmatrix} 2L \\ 2\ell_b \end{pmatrix}^{\frac{1}{2}} [(2L - 2\ell_b + 1)(2\ell_b + 1)(2L+1)]^{\frac{1}{2}}.$$

For instance the physical interactions: hexadecupole-monopole coupling (ℓ_a=4,ℓ_b=0), octupole-dipole coupling (ℓ_a=3,ℓ_b=1), and quadrupole-quadrupole coupling (ℓ_a=2,ℓ_b=2) all have an R^{-5} dependence.

CHAPTER II 3 LONG RANGE FORCES

In this chapter general formulas for the long range interaction between two systems (atoms or molecules) will be derived. As has been shown in ref. 7, these formulas form an asymptotic expansion (for large R) of the VB interaction energy.

From the outset we make the assumption that the monomer wave functions do not overlap. This has two immediate consequences:

- All intermolecular exchange integrals vanish, and so we are relieved of the difficult task of intermolecular antisymmetrization.
- The multipole expansion derived in the preceding chapter is applicable.

Let us choose two parallel space fixed coordinate systems A and B located at the centers of mass of the subsystems A and B, respectively. The vector $\vec{R}$ connects the origins. The particles (nuclei and electrons) of A are labelled by α. Their charge is denoted by q_α and their position vector with respect to the coordinate system at A by $\vec{r}_\alpha$. Similarly $\vec{r}_\beta$ gives the position with respect to the system at B of particle β, belonging to monomer B and having charge q_β. Monomer A may be rotated over a set of Euler angles $\omega_A = (\alpha_A, \beta_A, \gamma_A)$ and monomer B over $\omega_B = (\alpha_B, \beta_B, \gamma_B)$. Subsequently expressions will be derived for the first and second order interaction energy as a function of $\vec{R}$, ω_A and ω_B. It is apparent that these coordinates are redundant in a description of the complete interaction potential surface. For example just $\vec{R}$ and ω_B are sufficient.

Write the Hamiltonian of the system as:

$$H = H^A + H^B + V^{AB} \tag{1}$$

where:

$$V^{AB} = \sum_{\alpha,\beta} \frac{q_\alpha q_\beta}{r_{\alpha\beta}}. \tag{2}$$

The interaction can be expanded (see ch. II.2):

$$V^{AB} = \sum_{\alpha,\beta} q_\alpha q_\beta \sum_{\ell_A=0}^{\infty} \sum_{\ell_B=0}^{\infty} C_{\ell_A \ell_B} \sum_{\mu=-\lambda}^{+\lambda} (-1)^\mu Z_{\lambda,-\mu}(\vec{R}) \times \times [\vec{S}_{\ell_A}(\vec{r}_\alpha) \times \vec{S}_{\ell_B}(\vec{r}_\beta)]^{(\lambda)}_\mu \tag{3}$$

where:

$$\lambda \equiv \ell_A + \ell_B \tag{4}$$

$$c_{\ell_A \ell_B} = (-1)^{\ell_B} (4\pi)^{3/2} [(2\ell_A+1)(2\ell_B+1)(2\ell_A+2\ell_B+1)]^{-\frac{1}{2}} \begin{pmatrix} 2\ell_A + 2\ell_B \\ 2\ell_A \end{pmatrix}^{\frac{1}{2}}. \tag{5}$$

Let $\{|a\rangle\}$ be a complete and discrete set of eigenstates of H^A:

$$H^A|a\rangle = E_a|a\rangle, \qquad a = 0, 1, 2, \ldots, \infty \tag{6}$$

and similarly on B:

$$H^B|b\rangle = E_b|b\rangle, \qquad b = 0, 1, 2, \ldots, \infty. \tag{7}$$

Define a multipole (transition) moment on A:

$$\langle a|\vec{\ell}_A|a'\rangle \equiv \{\langle a|\ell_A, m_A|a'\rangle \; ; \; m_A = -\ell_A, \ldots, +\ell_A\} \tag{8}$$

where:

$$\langle a|\ell_A, m_A|a'\rangle \equiv \langle a|\sum_\alpha q_\alpha S_{\ell_A, m_A}(\vec{r}_\alpha)|a'\rangle. \tag{9}$$

In exactly the same way the multipole transition moment $\langle b|\vec{\ell}_B|b'\rangle$ is defined.

The transition moments are irreducible tensorial sets. Rotation of molecule A over ω_A gives the following mapping:

$$\langle a|\vec{\ell}_A|a'\rangle \rightarrow \langle a|\vec{\ell}_A \cdot \underline{D}^{(\ell_A)}(\omega_A^{-1})|a'\rangle \tag{10}$$

where:

$$\langle a|(\vec{\ell}_A \cdot \underline{D}^{(\ell_A)}(\omega_A^{-1}))_{m_A}|a'\rangle = \sum_{m_A'} D^{(\ell_A)}(\omega_A^{-1})_{m_A' m_A} \langle a|\ell_A, m_A'|a'\rangle. \tag{11}$$

Here $\underline{D}^{(\ell_A)}$ is the $(2\ell_A+1)$-dimensional irrep of SO(3). In an analogous manner one proceeds for the transition moments on B.

Up to second order perturbation the interaction energy ΔE^{AB} becomes:

$$\Delta E^{AB} = \Delta E^{(1)} + \Delta E^{(2)} \tag{12}$$

$$\Delta E^{(1)} = \langle 00|V^{AB}|00\rangle \tag{13}$$

$$\Delta E^{(2)} = \sum_{a,b=0}^{\infty}{}' (\Delta\varepsilon_{ab})^{-1} \langle 00|V^{AB}|ab\rangle\langle ab|V^{AB}|00\rangle \tag{14}$$

$$\Delta\varepsilon_{ab} = (E_{a=0} - E_a) + (E_{b=0} - E_b). \tag{15}$$

In eq. (13) and (14) a product ket $|a\rangle \otimes |b\rangle$ is written as $|ab\rangle$, with the first quantum number labelling a ket on A and the second a ket on B. The same holds for the bra.

The first order interaction energy can be written as an explicit function of $\vec{R}$, ω_A and ω_B, when we insert the multipole expansion (3) into (13) and use the rotation property (10):

$$\Delta E^{(1)}(\omega_A,\omega_B,\vec{R}) = \sum_{\ell_A=0}^{\infty}\sum_{\ell_B=0}^{\infty} C_{\ell_A\ell_B} \sum_{\mu=-\lambda}^{+\lambda} (-1)^{\mu} Z_{\lambda,-\mu}(\vec{R}) \times$$
$$\times \left[\langle 0 | \vec{\ell}_A \cdot \underline{D}^{(\ell_A)}(\omega_A^{-1}) | 0 \rangle \times \langle 0 | \vec{\ell}_B \cdot \underline{D}^{(\ell_B)}(\omega_B^{-1}) | 0 \rangle \right]^{(\lambda)}_{\mu} \qquad (16)$$

where $\lambda \equiv \ell_A + \ell_B$ and $C_{\ell_A\ell_B}$ is the coefficient defined in eq. (5). The expression between square brackets stands for the coupling of a 2^{ℓ_A}-pole on A with a 2^{ℓ_B}-pole on B to a resulting tensor of maximal degree $\lambda \equiv \ell_A + \ell_B$. The resulting tensor is coupled with an irregular harmonic to form an SO(3) scalar.

So, the terms with fixed ℓ_A and ℓ_B in this expansion of the first order energy represent a tensor interaction [14] between a (rotated) permanent moment on A and a (rotated) permanent moment on B. The total interaction is zero if one of the monomers has no permanent multipole moment, that is, if one of the subsystems is a neutral atom in an S-state. Interpreting the wave function squared as a classical charge distribution, eq. (16) can also be looked upon as the multipole expansion of the classical electrostatic interaction energy. So, in fact we have found a general expression for the tensor force which was considered by Keesom [3] up to quadrupole-quadrupole interaction ($\ell_A = \ell_B = 2$).

Because of the orthogonality relation in the full rotation group:

$$\frac{1}{8\pi^2} \int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} D^{(0)}(\alpha\beta\gamma)^*_{00}\, D^{(\ell)}(\alpha\beta\gamma)_{mm'}\, d\alpha \sin\beta\, d\beta\, d\gamma = \delta_{m0}\,\delta_{m'0}\,\delta_{\ell 0} \qquad (17)$$

(see ref. 11, eq. 4.6.1) and because $D^{(0)}(\alpha\beta\gamma)^*_{00} = 1$, it follows that the first order energy, averaged over all orientations of one of the subsystems, is zero. This is not to imply that powers of the first order interaction also vanish upon orientational averaging. The electrostatic energy does give a contribution to the second virial coefficient of an imperfect gas, but only in second and higher order of (1/kT).

The second order energy becomes after insertion of the multipole expansion (3) into eq. (14):

$$\Delta E^{(2)}(\vec{R}) = \sum_{\ell_A,\ell'_A} \sum_{\ell_B,\ell'_B} c_{\ell_A\ell_B} c_{\ell'_A\ell'_B} \sum_{a,b=0}^{\infty\prime} (\Delta\varepsilon_{ab})^{-1} Q(\ell_A,\ell_B;\ell'_A,\ell'_B) \quad (18)$$

$$Q(\ell_A,\ell_B;\ell'_A,\ell'_B) = \sum_{\mu=-\lambda}^{+\lambda} \sum_{\mu'=-\lambda'}^{+\lambda'} (-1)^{\mu+\mu'} Z_{\lambda,-\mu}(\vec{R})\, Z_{\lambda',-\mu'}(\vec{R})\, T_{\lambda,\mu}\, T'_{\lambda',\mu'} \quad (19)$$

where:

$$T_{\lambda,\mu} \equiv [\langle 0|\vec{\ell}_A|a\rangle \times \langle 0|\vec{\ell}_B|b\rangle]^{(\lambda)}_{\mu} \quad (20)$$

$$T'_{\lambda',\mu'} \equiv [\langle a|\vec{\ell}'_A|0\rangle \times \langle b|\vec{\ell}'_B|0\rangle]^{(\lambda')}_{\mu'}. \quad (21)$$

This expression for the second order energy has two disadvantages: the $\vec{R}$-dependence is given as a product of two irregular solid harmonics, which is inconvenient; furthermore this expression becomes a rather complicated function of ω_A and ω_B after substitution of rotation matrices in the manner of the first order energy (16). Therefore we will rewrite this expression, so that all interactions are given in terms of irreducible tensorial products.

Consider the quantity $Q(\ell_A,\ell_B;\ell'_A,\ell'_B)$ defined in equation (19). Write:

$$Z_{\lambda,-\mu}\, Z_{\lambda',-\mu'} = \sum_{L,M} (L,M|\lambda,-\mu;\lambda',-\mu')\, [\vec{Z}_\lambda \times \vec{Z}_{\lambda'}]^{(L)}_M \quad (22)$$

$$T_{\lambda,\mu}\, T'_{\lambda',\mu'} = \sum_{L',M'} (L',M'|\lambda,\mu;\lambda',\mu')[\vec{T}_\lambda \times \vec{T}'_{\lambda'}]^{(L')}_{M'}. \quad (23)$$

Use:

$$(L,M|\lambda,-\mu;\lambda',-\mu') = (-1)^{\lambda+\lambda'+L} (L,-M|\lambda,\mu;\lambda',\mu'),$$

apply the orthogonality of C-G coefficients and we get:

$$Q(\ell_A,\ell_B;\ell'_A,\ell'_B) = (-1)^{\lambda+\lambda'} \sum_L (-1)^L \sum_M (-1)^M [\vec{Z}_\lambda \times \vec{Z}_{\lambda'}]^{(L)}_{-M} [\vec{T}_\lambda \times \vec{T}'_{\lambda'}]^{(L)}_M. \quad (24)$$

From the Gaunt series (ref. 11, eq. 5.1.6) one derives easily:

$$[\vec{Z}_\lambda \times \vec{Z}_{\lambda'}]^{(L)}_{-M} = \left[\frac{(2\lambda+1)(2\lambda'+1)}{4\pi(2L+1)}\right]^{\frac{1}{2}} (\lambda,0;\lambda',0|L,0)\, R^{-(\lambda+\lambda'+2)}\, Y_{L,-M}(\hat{R}). \quad (25)$$

The product $[\vec{T}_\lambda \times \vec{T}'_{\lambda'}]^{(L)}_M$ is in fact a fourfold irreducible product (see (20) and (21)). We can recouple this product:

$$[[\langle 0|\vec{\ell}_A|a\rangle \times \langle 0|\vec{\ell}_B|b\rangle]^{(\lambda)} \times [\langle a|\vec{\ell}'_A|0\rangle \times \langle b|\vec{\ell}'_B|0\rangle]^{(\lambda')}]^{(L)}_M =$$

$$= \sum_{L_A}\sum_{L_B} (\ell_A\ell'_A)L_A,(\ell_B\ell'_B)L_B,L|(\ell_A\ell_B)\lambda,(\ell'_A\ell'_B)\lambda',L) \times \tag{26}$$

$$\times [\vec{M}^{oa}_{(\ell_A\ell'_A)L_A} \times \vec{M}^{ob}_{(\ell_B\ell'_B)L_B}]^{(L)}_M .$$

Here:

$$L_A = |\ell_A - \ell'_A|, \ldots, \ell_A + \ell'_A$$

$$L_B = |\ell_B - \ell'_B|, \ldots, \ell_B + \ell'_B$$

$$\vec{M}^{oa}_{(\ell_A\ell'_A)L_A} \equiv [\langle 0|\vec{\ell}_A|a\rangle \times \langle a|\vec{\ell}'_A|0\rangle]^{(L_A)} \tag{27}$$

$$\vec{M}^{ob}_{(\ell_B\ell'_B)L_B} \equiv [\langle 0|\vec{\ell}_B|b\rangle \times \langle b|\vec{\ell}'_B|0\rangle]^{(L_B)} . \tag{28}$$

The recoupling coefficient is (ref. 11, eq. 6.4.2):

$$(\ell_A\ell'_A)L_A,(\ell_B\ell'_B)L_B,L|(\ell_A\ell_B)\lambda,(\ell'_A\ell'_B)\lambda',L) =$$

$$= [(2L_A+1)(2L_B+1)(2\lambda+1)(2\lambda'+1)]^{\frac{1}{2}} \begin{Bmatrix} \ell_A & \ell'_A & L_A \\ \ell_B & \ell'_B & L_B \\ \lambda & \lambda' & L \end{Bmatrix} \tag{29}$$

The expression between curly brackets is the Wigner 9j-symbol. Substitution of (25), (26) and (29) into (24) yields:

$$Q(\ell_A,\ell_B;\ell'_A\ell'_B) = (-1)^{\lambda+\lambda'}\sqrt{\frac{1}{4\pi}}\, R^{-(\lambda+\lambda'+2)} \sum_{L_A}\sum_{L_B}\sum_{L}\sum_{M} (-1)^{L+M} \times$$

$$\times d(\ell_A\ell'_A L_A;\ell_B\ell'_B L_B;L)\; Y_{L,-M}(\hat{R})\; [\vec{M}^{oa}_{(\ell_A\ell'_A)L_A} \times \vec{M}^{ob}_{(\ell_B\ell'_B)L_B}]^{(L)}_M \tag{30}$$

where:

$$d(\ell_A\ell'_A L_A;\ell_B\ell'_B L_B;L) = (2\lambda+1)(2\lambda'+1)\left[\frac{(2L_A+1)(2L_B+1)}{(2L+1)}\right]^{\frac{1}{2}} \times$$

$$\times (\lambda,0;\lambda'0|L,0) \begin{Bmatrix} \ell_A & \ell'_A & L_A \\ \ell_B & \ell'_B & L_B \\ \lambda & \lambda' & L \end{Bmatrix} \tag{31}$$

Substituting (30) into (18) and further using the fact that the coupled moments, defined in (27) and (28), are irreducible sets, we arrive finally at:

$$\Delta E^{(2)}(\omega_A,\omega_B,\vec{R}) = (4\pi)^{5/2} \sum_{\ell_A,\ell'_A} \sum_{\ell_B,\ell'_B} (-1)^{\ell_A+\ell'_A} \times$$

$$\times A(\ell_A,\ell'_A,\ell_B,\ell'_B)\, R^{-(\ell_A+\ell'_A+\ell_B+\ell'_B+2)} \times$$

$$\times \sum_{L_A} \sum_{L_B} \sum_{L} B((\ell_A\ell'_A)L_A,(\ell_B\ell'_B)L_B,L) \times \tag{32}$$

$$\times \sum_{a,b=0}^{\infty}{}' (\Delta\epsilon_{ab})^{-1} \sum_{M} (-1)^M Y_{L,-M}(\bar{R}) \times$$

$$\times[(\vec{M}^{oa}_{(\ell_A\ell'_A)L_A} \cdot \underline{D}^{(L_A)}(\omega_A^{-1})) \times (\vec{M}^{ob}_{(\ell_B\ell'_B)L_B} \cdot \underline{D}^{(L_B)}(\omega_B^{-1}))]^{(L)}_{M}$$

where:

$$\ell_A,\ell'_A = 0,\ 1,\ 2,\ \ldots,\ \infty$$

$$\ell_B,\ell'_B = 0,\ 1,\ 2,\ \ldots,\ \infty$$

$$L_A = |\ell_A - \ell'_A|,\ |\ell_A - \ell'_A| + 1,\ \ldots,\ \ell_A + \ell'_A$$

$$L_B = |\ell_B - \ell'_B|,\ |\ell_B - \ell'_B| + 1,\ \ldots,\ \ell_B + \ell'_B$$

$$L = |L_A - L_B|,\ |L_A - L_B| + 1,\ \ldots,\ L_A + L_B$$

$$M = -L,\ -L+1,\ \ldots,\ L$$

and:

$$A(\ell_A,\ell'_A,\ell_B,\ell'_B) = \left[\frac{(2\ell_A+2\ell_B+1)!\ (2\ell'_A+2\ell'_B+1)!}{(2\ell_A+1)!\ (2\ell_B+1)!\ (2\ell'_A+1)!\ (2\ell'_B+1)!}\right]^{\frac{1}{2}} \tag{33}$$

$$B((\ell_A\ell'_A)L_A,(\ell_B\ell'_B)L_B,L) = [(2L_A+1)(2L_B+1)]^{\frac{1}{2}} \begin{pmatrix} \ell_A+\ell_B & \ell'_A+\ell'_B & L \\ 0 & 0 & 0 \end{pmatrix} \times$$

$$\times \begin{Bmatrix} \ell_A & \ell'_A & L_A \\ \ell_B & \ell'_B & L_B \\ \ell_A+\ell_B & \ell'_A+\ell'_B & L \end{Bmatrix} \tag{34}$$

A term with fixed L_A and L_B in the second order interaction (32) represents the coupling of the tensor $\vec{M}^{oa}_{(\ell_A \ell'_A) L_A}$ centered on A with a tensor $\vec{M}^{ob}_{(\ell_B \ell'_B) L_B}$ centered on B. Hence the second order interaction is, just like the first order interaction, a tensor interaction [14], but contrary to the case of the first order interaction the two monomer tensors do not only couple to their maximal L-value ($L = L_A + L_B$), but to all their possible lower L-values as well. Note that the monomer tensor $\vec{M}^{oa}_{(\ell_A \ell'_A) L_A}$, which is of order L_A, originates from the coupling of a 2^{ℓ_A}-pole on A with a $2^{\ell'_A}$-pole on A. In the same way the tensor $\vec{M}^{ob}_{(\ell_B \ell'_B) L_B}$ arises on monomer B from the coupling of two multipoles on B.

It is customary to distinguish three different contributions to the second order energy. To that end one splits the sum over the excited states of the dimer thus:

$$\sum_{a,b=0}^{\infty}{}' = \sum_{b=1}^{\infty} + \sum_{a=1}^{\infty} + \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \tag{35}$$

The first term in the right hand side gives rise to interactions of the type:

$$\Delta E^{A \to B}_{(ind)} \propto R^{-(\ell_A + \ell'_A + \ell_B + \ell'_B + 2)} \sum_M (-1)^M Y_{L,-M}(\hat{R}) \times$$
$$\times \left[\vec{M}^{oa}_{(\ell_A \ell'_A) L_A} \times \vec{P}_{(\ell_B \ell'_B) L_B} \right]^{(L)}_M . \tag{36}$$

Here:

$$\vec{P}_{(\ell_B \ell'_B) L_B} \equiv \sum_{b=1}^{\infty} (\Delta\varepsilon_{ob})^{-1} \vec{M}^{ob}_{(\ell_B \ell'_B) L_B} \tag{37}$$

is an irreducible polarizability tensor giving a 2^{ℓ_B}-pole and a $2^{\ell'_B}$-pole on B, which in eq. (36) are interacting with a permanent 2^{ℓ_A}-pole and $2^{\ell'_A}$-pole on A. The interaction (36) has a classical counterpart in Debije's induction force [4]. Debije considered the special case $\ell_A = \ell'_A = 2$, $\ell_B = \ell'_B = 1$, that is, a permanent quadrupole on A inducing a dipole on B. This interaction has an R^{-8} dependence.

The second term in (35) gives a sum of interactions, all arising from the induction of multipoles on A by permanent moments on B.

The last term does not allow a classical interpretation, it gives

London's dispersion force [5]. London considered the special case $\ell_A = \ell_A' = \ell_B = \ell_B' = 1$, the induced-dipole-induced-dipole interaction. This interaction has an R^{-6} dependence.

When we average the second order energy (32) over all orientations of the monomers:

$$\langle \Delta E^{(2)}(\vec{R}) \rangle = \int_{\omega_A} \int_{\omega_B} \Delta E^{(2)}(\omega_A, \omega_B, \vec{R})\, d\omega_A\, d\omega_B \tag{38}$$

and use (17), we obtain only a contribution for $L_A = 0$ and $L_B = 0$. And since:

$$B((\ell_A \ell_A')0,(\ell_B \ell_B')0,0) = \delta_{\ell_A \ell_A'}\, \delta_{\ell_B \ell_B'}\, (-1)^{\ell_A+\ell_B}\, (2\ell_A + 2\ell_B + 1)^{-\frac{1}{2}} \tag{39}$$

we get:

$$\begin{aligned}\langle \Delta E^{(2)}(R) \rangle = (4\pi)^2 \sum_{\ell_A,\ell_B} (-1)^{\ell_A+\ell_B} [(2\ell_A + 1)(2\ell_B + 1)]^{-3/2} \times \\ \times \binom{2\ell_A + 2\ell_B}{2\ell_A} R^{-2(\ell_A+\ell_B+1)} \sum_{a,b}{}' (\Delta\epsilon_{ab})^{-1} \times \\ \times \sum_{m_A} (-1)^{m_A} \langle 0|\ell_A, m_A|a\rangle \langle a|\ell_A, -m_A|0\rangle \times \\ \times \sum_{m_B} (-1)^{m_B} \langle 0|\ell_B, m_B|b\rangle \langle b|\ell_B, -m_B|0\rangle .\end{aligned} \tag{40}$$

The matrix elements arising in this expression are defined in (9). For instance the average dipole-dipole interaction is.

$$\begin{aligned}\langle \Delta E^{(2)}_{d-d}(R) \rangle = 2/3\, R^{-6} \sum_{a,b}{}' (\Delta\epsilon_{ab})^{-1} (\langle 0|\vec{r}_A|a\rangle \cdot \langle a|\vec{r}_A|0\rangle) \times \\ \times (\langle 0|\vec{r}_B|b\rangle \cdot \langle b|\vec{r}_B|0\rangle)\end{aligned} \tag{41}$$

where

$$\langle 0|\vec{r}_A|a\rangle \cdot \langle a|\vec{r}_A|0\rangle = |\langle 0|\textstyle\sum_\alpha x_\alpha|a\rangle|^2 + |\langle 0|\sum_\alpha y_\alpha|a\rangle|^2 + |\langle 0|\sum_\alpha z_\alpha|a\rangle|^2 \tag{42}$$

and a similar definition holds for $\langle 0|\vec{r}_B|b\rangle \cdot \langle b|\vec{r}_B|0\rangle$. This same formula (41) has been derived by London [5] for the case of two interacting atoms.

London averaged by summing over the degenerate states of the atoms. Note, however, that equation (40) is completely general, no assumption whatever has been made about the symmetry of H^A and H^B, so the formula is valid for molecules of arbitrary symmetry.

Notes.

1. A formula similar to (40) has been derived by Riera and Meath [15]. These authors take the orientational average of the second order energy given in eq. (18). Since formula (18) is not simplified to the utmost with respect to its transformation properties, the final formula of ref. 15 is not as simple as our expression (40). Furthermore the result of Riera and Meath contains a small error; it seems that the authors have not divided by the volume of SO(3) and hence they present an unnormalized averaged interaction energy.
2. Contrary to the electrostatic (first order) interaction, neither the induction energy nor the dispersion energy vanishes upon orientational averaging. The leading contribution to the second virial coefficient arising from these forces is therefore proportional to 1/kT and hence is in general much larger than the electrostatic contribution, although the (non-averaged) electrostatic interaction is often larger than the (non-averaged) induction and (non-averaged) dispersion interaction.

The second order interaction energy computed above is always attractive, as can be seen from eq. (14). Nonetheless it is experimentally known that two closed-shell molecules repel each other at short distances, and so there is something lacking in the theory. Now, one of the basic assumptions we made in the beginning of this chapter was that intermolecular antisymmetrization is unnecessary, but, as we have already discussed in ch. II.1, intermolecular exchange causes repulsion and so we miss this effect. One can remedy this fault by including the intermolecular antisymmetrizer into the perturbation formulas. This approach has already been pursued by Eisenschitz and London [1]; more recently their results have been recast into a wave operator formalism by Van der Avoird [16]. Many other workers have also attempted to derive a practical exchange-perturbation theory [17], and several authors have applied such a theory to calculations on simple systems [18, 19]. The general experience is

that a second order energy which includes exchange is difficult to compute for any but the simplest systems. Moreover the practical exchange-perturbation calculations usually proceed by a variational approximation. Therefore, we decided rather to use the variational principle from the start.

The first of the following papers is based on a semi-empirical SCF method. This approach has the great advantage that relatively large complexes can be handled. However, as already has been discussed in the introduction (ch. II.1), the outcome of these calculations was not very promising. Therefore we turned to the ab initio multi-structure VB method. The paper on He_2 must be considered as a first test on this formalism; the work on the ethylene dimer is an application of the VB method to a chemically more interesting complex. The paper on $He-H_2$ has been an attempt to obtain an intermolecular potential of such quality that it could be of use in molecular beam scattering experiments. This attempt has proved to be successful [20].

REFERENCES PART II

1. R. Eisenschitz, F. London,
   Z. Physik 60, 491 (1930).

2. W. Heitler,
   Elementary Wave Mechanics, 2nd ed.
   Oxford Univ. Press, London, 1956.

3. W.H. Keesom,
   Phys. Zeitschr. 22, 129 (1921).

4. P. Debije,
   Phys. Zeitschr. 21, 178 (1920).

5. F. London,
   Z. Physik 63, 245 (1930).

6. F. London,
   Z. Physik.Chem. (B) 11, 222 (1931).

7. P.E.S. Wormer, A. van der Avoird,
   J. Chem. Phys. 62, 3326 (1975).
   (Reprinted in this thesis).

8. A. van der Avoird, D.J.M. Fassaert, P.E.S. Wormer,
   Quantumtheoretische Methoden voor Atomen en Moleculen.
   Lecture Notes, University of Nijmegen, 1975.

9. A.D. Buckingham,
   Adv. Chem. Phys. 12, 107 (1967).

10. M.E. Rose,
    J. Math. and Phys. 37, 215 (1958).

11. A.R. Edmonds,
    Angular Momentum in Quantum Mechanics.
    Princeton Univ. Press, Princeton, 1957.

12. D.M. Brink, G.R. Satchler,
    Angular Momentum, 2nd ed.
    Oxford Univ. Press, London, 1968.

13. B.R. Judd,
    Operator Techniques in Atomic Spectroscopy.
    McGraw-Hill, New York, 1963.

14. A. de-Shalit, I. Talmi,
    Nuclear Shell Theory.
    Academic Press, New York, 1963.

15. A. Riera, W.J. Meath,
    Int. J. Quant. Chem. 7, 959 (1973).

16. A. van der Avoird,
J. Chem. Phys. 47, 3649 (1967).

17. D.M. Chipman, J.D. Bowman, J.O. Hirschfelder,
J. Chem. Phys. 59, 2830 (1973),
and references therein.

18. A. van der Avoird,
Thesis, Technical University of Eindhoven, 1968.

19. D.M. Chipman, J.O. Hirschfelder,
J. Chem. Phys. 59, 2838 (1973).

20. L. Zandee, J. Verberne, J. Reuss,
Submitted for publication in Chem. Phys. Lett.

The TCNE-Benzene Complex: A CNDO Approach*

D. B. CHESNUT and PAUL E. S. WORMER**
Paul M Gross Chemical Laboratory, Duke University, Durham, North Carolina 27706

Received October 9, 1970

* Supported in part by NASA University Sustaining Grant NGR 34-001-005 and National Science Foundation Grant GP-8298

** Present address Department of Theoretical Chemistry, University of Nijmegen, The Netherlands

CNDO/2 calculations on the TCNE-benzene complex are reported A stable complex is found which exhibits a relatively large stabilization energy (0 2 a u) at a short interplanar separation (1 75 Å), the binding apparently arises solely through charge transfer Mulliken population analyses were performed by reinterpreting the CNDO orbitals as Lowdin orbitals Sample calculations on small organic molecules and first row diatomics indicate the procedure to be satisfactory It is shown that generally only overlap populations that are summed over the orbitals of the atoms in question reflect the symmetry of the molecule

Die Ergebnisse von CNDO/2-Rechnungen an Tetracyanaoathylen-Benzol-Komplexen werden mitgeteilt. Es wird ein stabiler Komplex gefunden, der eine relativ große Stabilisierungsenergie (0,2 A E) bei geringem Abstand (1,75 Å) der Molekülebenen besitzt, die Bindung entsteht anscheinend nur durch Ladungsubertragung Eine Populationsanalyse nach Mulliken wurde mit Hilfe der Interpretation der CNDO-Orbitale als Lowdin-Orbitale durchgefuhrt Berechnungen an Beispielen wie kleinen organischen Molekulen und zweiatomigen Molekulen aus Elementen der ersten Reihe zeigen, daß die Methode befriedigende Ergebnisse liefert Es wird gezeigt, daß im allgemeinen nur die Überlappungs-Populationen, die uber die Orbitale der betrachteten Atome summiert werden, die Symmetrie des Molekuls widerspiegeln

Calculs CNDO/2 sur le complexe TCNE-benzene Un complexe stable apparaît pour une séparation interplan courte (1,75 Å) avec une energie de stabilisation relativement forte (0,2 u a); la liaison provient apparemment du seul transfert de charge. Une analyse de population de Mulliken a été effectuée en réinterprétant les orbitales CNDO comme orbitales de Löwdin Des calculs échantillonés sur de petites molecules organiques et des molécules diatomiques de la première ligne montrent que le procédé s'avère satisfaisant On montre qu'en général, seules les populations de recouvrement sommées sur les orbitales des atomes en question reflètent la symetrie de la molécule

Introduction

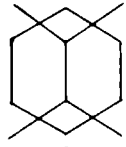
π-π molecular complexes pose an interesting study in bonding. The question of the prime effect in stabilizing such systems is not yet clear due mainly to the difficulty in adequately treating such large systems. For some time it was generally felt that charge transfer stabilization was mainly responsible for bonding, but spectroscopic measurements on some TCNE (tetracyanoethylene) complexes do not seem to support this idea [1]. Studies employing a π-only extended Hückel approach [2] have given indications of reproducing relative molecular geometries at chosen experimental interplanar intermolecular separations but in general fail to predict absolute binding. The advent of approximate self-consistent field

treatments introduced by Pople and coworkers [3] allows a more readily justifiable treatment of large systems and avoids many of the pitfalls of one-electron methods For example, several recent calculations [4, 5] using this method on hydrogen bonded systems have yielded reasonable results. In the present paper we present CNDO/2 calculations of the TCNE-benzene complex for a restricted region of relative intermolecular geometry

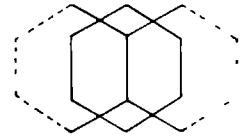
A basic purpose for initiating such a calculation is the elucidation of those factors which dictate the observed relative intermolecular geometry characteristic of the majority of both charge transfer and charge resonance complexes Unfortunately, the current calculations do not reveal any particular simplifying feature, on the contrary, the issue is clouded a bit by our results which indicate small and negative intermolecular overlap populations and a high degree of charge transfer It is worthwhile to discuss these results in light of the general problems above. Too, since no previous CNDO calculations on systems of this kind have been performed the present calculations may be taken as a model study of the CNDO approach to a π-bonded bimolecular system.

Choice of Model

The dominant structural fragment seen in most π-π molecular complexes consists of a bonds-over-bonds arrangement shown below,

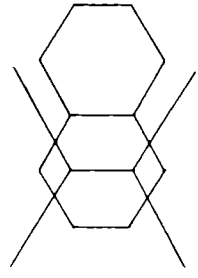


a



b

where the benzenoid moiety projects onto an ethylene-like fragment arising either from a real ethylene group (case a, as, for example, the central portion of TCNE itself) or from a fused benzenoid system (case b) Examples of such structures can be seen in Chesnut and Moseley's paper [2] and from the review paper of Prout and Wright [6] We wished to carry out calculations on a system that exhibits such a characteristic projected structure and at the same time was small enough to minimize computer time The available program is limited to 80 orbitals which severely restricts our selection The TCNE-benzene system represents a 70-orbital problem and, although its structure is not known, one would expect a behavior similar to the general class of materials Indeed, the TCNE-naphthalene crystal structure is known [7], exhibiting the projection



An interesting aspect of the TCNE-naphthalene structure is that as the TCNE and naphthalene molecules stack alternately with planes parallel to one another the TCNE molecules project onto one ring of the naphthalene below and the opposite ring of the naphthalene above.

Since we are interested in general trends we have concentrated on the above type of projected structure in the TCNE-benzene calculations. We have performed a few calculations on other geometries and, as discussed later, find that a rotated form of the complex exhibits greater binding according to the computational method we have employed. Aside from questioning the general integrity of the calculational approach, this raises the well-known problem of comparing calculations on an isolated molecular system (a gas-phase calculation) with experimental structural data obtained from studies of crystals.

Calculations and Method of Population Analysis

The calculations were performed on the IBM 360/75 computer at the Triangle Universities Computer Center, Research Triangle Park, North Carolina, using a modified version of the semiempirical self-consistent field CNINDO program supplied by the Quantum Chemistry Program Exchange, University of Indiana, Bloomington, Indiana. The semiempirical parameters used were the CNDO/2 parameters given by Pople *et al.* [3]. The internal geometries of TCNE and benzene were held fixed throughout the calculations. For TCNE, 120° angles plus the distances reported in Bekoe and Trueblood's crystal structure investigation [8] were used; for benzene, standard hexagonal symmetry was assumed with C–C and C–H distances of 1.40 Å and 1.08 Å, respectively. The molecular planes were always maintained parallel to each other.

Two modifications to the basic program were made. In its original form CNINDO showed diverging SCF cycles for TCNE and TCNE-benzene. To circumvent this difficulty the program was modified so that in the $(n+1)^{\text{th}}$ iterative SCF cycle the entered density matrix (charge and bond-order matrix), $\boldsymbol{P}^{(n)}$, used was taken to be

$$\boldsymbol{P}^{(n)} = \boldsymbol{P}^{(n-1)} + \delta(\boldsymbol{P}^{(n)'} - \boldsymbol{P}^{(n-1)}),$$

where $\boldsymbol{P}^{(n-1)}$ is the density matrix entering the n^{th} cycle, $\boldsymbol{P}^{(n)'}$ is the density matrix *exiting* the n^{th} cycle, and δ is a damping factor. A value of 0.7 for δ was found to yield optimum convergence for TCNE (convergence in 14 cycles) and proved satisfactory for the TCNE-benzene complex and another large cyanocarbon, TCNQ (tetracyanoquinodimethane). An average calculation on the TCNE-benzene complex took about 12.5 minutes.

The second modification involved the addition of a subroutine to perform a Mulliken population analysis [9] in terms of a transformed basis. One of the main points of interest of this work was the question of whether or not one can detect intermolecular bonding in the complex. One such measure has been simply to use the non-diagonal elements of the charge and bond-order matrix, the bond orders [10]. We are interested in the total measure of bonding between any two atoms and would thus want to sum individual bond orders of the various atom-

localized orbitals. This procedure, however, fails in general for an arbitrary choice of local atomic coordinate systems (to which the localized AO's are attached); physically equivalent (symmetry related) bonds can have different total summed bond orders (See Appendix). A simple example is to note that the π-bond order between a pair of P_z-π orbitals in an aromatic system is changed in sign by inverting one of the local z-axes. The fact that the use of simple bond orders is justified in π-electron calculations depends upon the standard choice of parallel local axes normal to the π-plane.

A way of avoiding this problem is to use the sum of Mulliken orbital overlap populations. We show in the Appendix that these summed populations reflect the symmetry of the molecules. The CNDO orbitals, however, are orthogonal in the zero-differential-overlap approximation so that there are no non-vanishing overlap populations. We choose, however, to consider the CNDO orbital as Löwdin orbitals [11], that is, semi-localized orbitals which resemble localized atomic orbitals as closely as possible [12] and yet are mutually orthogonal. Thus, the CNDO orbital set $\{\phi_n\}$ is taken to be related to a localized, non-orthogonal set $\{\phi'_n\}$ by the transformation.

$$\boldsymbol{\phi}' = \boldsymbol{\phi} \boldsymbol{S}^{1/2},$$

where $\boldsymbol{S}$ is the overlap matrix in the basis $\{\phi'_n\}$ and was constructed from the overlap integrals of the Slater orbitals used in the CNDO program. Expressing our CNDO wave functions in the $\{\phi'_n\}$ basis then allows one to perform a

Table 1 *Comparison of bond overlap populations*

Umbracketed values Present work, bracketed values Ref [14], part A and Ref [15], part B				
Molecule	Bond	σ	π	Total
A Li_2	--	0.872 (0.962)	0.0 (0.0)	0.872 (0.962)
N_2	-	0.431 (0.220)	0.880 (0.880)	1.311 (1.100)
F_2	-	0.256 (0.297)	−0.017 (−0.001)	0.239 (0.296)
LiF	--	0.155 (0.190)	0.322 (0.208)	0.477 (0.398)
LiH	--	0.757 (0.691)	0.0 (0.0)	0.757 (0.691)
B benzene	C−C	0.876 (0.82)	0.240 (0.240)	1.116 (1.06)
	C−H		--	0.768 (0.78)
ethylene	C=C	0.916 (0.791)	0.428 (0.420)	1.344 (1.211)
	C−H			0.778 (0.812)
ethane	C-C		-	0.856 (0.710)
	C−H	--	-	0.766 (0.785)
acetylene	C−C	0.390 (0.544)	0.920 (0.918)	1.310 (1.462)
	C−H		-	0.810 (0.800)
cis-butadiene	C=C	0.910 (0.77)	0.412 (0.39)	1.322 (1.16)
	C-C	0.832 (0.81)	0.050 (0.09)	0.882 (0.92)
	C−H	--	--	0.764, 0.780, 0.798 (0.78)
HCN	C≡N	0.390 (0.544)	0.920 (0.918)	1.310 (1.462)
	C--H	-		0.810 (0.800)
TCNE	C=C	0.818	0.376	1.194
	C−C	0.844	0.052	0.896
	C≡N	0.824[a]	0.448[b]	1.272

[a] Includes the equivalent of one (in plane) π-bond

[b] The "pure" π-bond

Mulliken population analysis in terms of net atomic charges and total bond overlap populations and this is the procedure we have followed. Recently, Brown and Roby [13] have shown that this interpretation of zero-differential-overlap theories can be justified from first principles.

To test this procedure we calculated total bond orders for a series of first row diatomics and small organic molecules and compared our results with literature values. The agreement is quite adequate; the comparisons for some representative systems are shown in Table 1, where the present results are compared with the calculations of Fraga and Ransil [14] for the diatomics and with the calculations of Lipscomb and coworkers [15] for the polyatomics. We used the same diatomic distances as Fraga and Ransil; for the organic systems we employed values from the "Tables of Interatomic Distance and Configurations in Molecules and Ions" (The Chemical Society, London, 1965) which differ very slightly in a few instances with the values of Newton, Boer, and Lipscomb.

Results and Discussion

Although no experimental gas phase data are available for the stabilization energy, the dipole moment and the intermolecular of the TCNE-benzene complex, one can get a rough idea of the order of magnitude of these quantities by comparison with similar complexes. In Table 2 some illustrative experimental data for complexes in which TCNE acts as the acceptor are presented.

To date no calculations, except those of Chesnut and Moseley [2] on the TCNE-anthracene complex, have been performed that vary the interplanar distance. Lippert, Hanna, and Trotter [18] take the fixed interplanar distance 3.50 Å and calculate by means of Morrell's perturbation theory [16] a dipole moment of 0.54 D and a stabilization energy of −6.21 kcal/mole. Their method is especially devised for charge transfer complexes and they use some parameters that are determined for the case under consideration. Herndon and Feuer [20] take a fixed distance as well (3.20 Å). They apply a perturbational method using as first-order functions for the separate molecules CNDO/2 functions and find a −3.54 kcal/mole stabilization energy. Mantione [21] calculated Van der Waals – London interactions in TCNE complexes with *p*-xylene, *o*-xylene, mesitylene and durene. His results agree remarkably well with the experimental results. Again the distance was fixed (3.30 Å).

The current calculations represent an attempt at a somewhat more general approach in which the relative geometry is determined by the calculation. Even

Table 2 *Illustrative data for some* TCNE *complexes*

Donor	ΔE (kcal mole)	μ (Debye)	R (Å)
benzene	− 3.34[a]	-	- -
durene	−10.1[b]	1.26[c]	—
naphthalene	− 4.06[d]	1.28[e]	3.30[f]

[a] In CCl_4 (Ref [16])
[b] In gas phase (Ref [17])
[c] In CCl_4 (Ref [16]).
[d] In CCl_4 (Ref [16])
[e] In CCl_4 (Ref [15])
[f] In solid state (Ref [7])

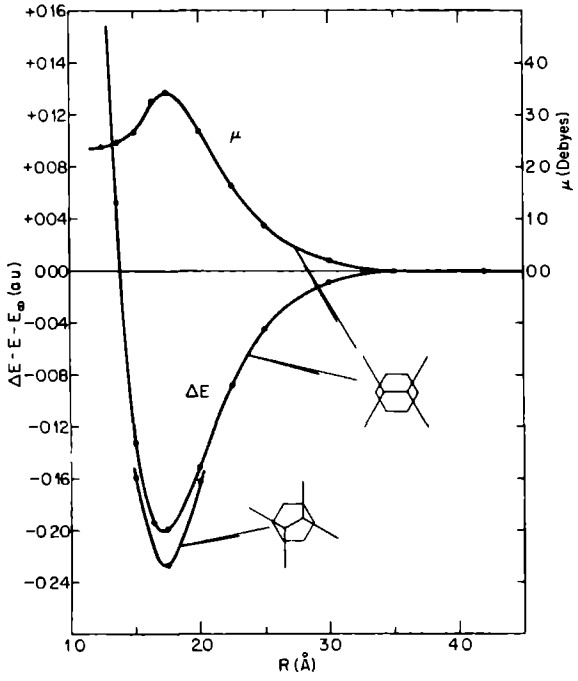


Fig 1 $\Delta E = E - E_{\infty}$ (in a u) as a function of the interplanar separation, R (in Å), for the relative geometries studied μ, the dipole moment (in Debyes) is also shown for the non-rotated structure The structure projections are approximately to scale and illustrate the two relative geometries studied

so, finite computer time and funds necessitated a limitation on the possible geometries to be studied Most of the calculations refer to that situation in which the two molecules are brought together directly on top of each other with the projected TCNE double bond direction colinear with the 1,4-carbon atoms on the benzene ring (the non-rotated form), a few points for the geometry in which the TCNE was rotated by 30 about the axis normal to its plane (the rotated forms) were also calculated The non-rotated and 30° rotated geometries are shown in projection as part of Fig 1 One must note that the optimum procedure in calculations of this kind would be to vary all distances to find the minimum energy, this would require very extensive calculations for the present case, calculations which we don't believe are warranted Thus, in all our calculations the *internal* geometries of the component molecules are held fixed as the distance between them, R, was varied

Fig 1 shows the calculated energy versus the interplanar separation for both the rotated and non-rotated structures, in Table 3 the electronic and nuclear energy components are listed along with the total energy While the non-rotated structure is the one for which most calculations were carried out, the rotated form (30) at $R = 1\,75$ Å actually exhibits a lower minimum ($E_{30} < E_{15} < E_0$ at 1 75 Å) Relative to one's *a priori* prejudices, the energy curve shows too deep a minimum at too short an interplanar separation with an extremely steep variation

Table 3 *Relative energies (in a u) as a function of R*

E_N = nuclear repulsions, E_{el} = electronic energy = $E - E_N$, E = total energy = $E_N + E_{el}$

R (Å)	$E_N - E_{N\infty}$	$E_{el} - E_{el\infty}$	$\Delta E = E - E_\infty$
1 25	271 401	−270 925	0 476
1 375	262 258	−262 206	0 052
1 50	253 567	−253 700	−0 133
1 64	244 349	−244 543	−0 194
1 75[a]	237 471	−237 670	−0 199
1 75[b]	237 417	−237 629	−0 212
1 75[c]	237 358	−237 586	−0 228
2 00	222 948	−223 099	−0 151
2 25	209 827	−209 915	−0 088
2 50	197 945	−197 990	−0 045
3 00	177 339	−177 348	−0 009
3 50	160 180	−160 181	−0 001

[a] 0° rotation — [b] 15° rotation [c] 30° rotation

of the energy at both small and large values or R Whereas stabilization energies of the order of 2 to 5 kcal/mole might be expected, an energy minimum here is calculated to be approximately 125 kcal/mole Whereas interplanar separations *in crystals* are from 3.2 to 3.5 Å, the minimum here occurs at approximately 1.75 Å. The energy curve at large R approaches an R^{-14} behavior with no hint of the London R^{-6} dispersion force behavior. The absence of the London forces in our results need not be surprising since it is generally held that these forces cannot be reproduced in a self-consistent field calculation without configuration interaction.

The large value of ΔE, the energy of stabilization, is somewhat characteristic of the CNDO method which tends to overemphasize binding This property may well also explain why the minimum in ΔE occurs at too short a distance To improve usefulness of these calculations a detailed study of what parameters or integrals led to this effect should be made and the necessary reparameterization be made As alluded to above, the importance of configuration interaction must also be ascertained The lowering of ΔE_{min} that occurs in going to the rotated complex occurs through a lowering of the nuclear repulsion terms which overcome an *increase* in the electronic energy (see Table 3) It is clear that the detailed geometry may well be very sensitive to the specific input parameters in systems as large and complicated as the present one Jesaitis and Streitwieser [22] point out that because the CNDO method uses *s*-orbitals to calculate the coulomb repulsion, the electron-electron repulsive energy may be underestimated in systems where the charge can be highly delocalized This effect may be quite significant in the present case since transfer of charge from benzene to TCNE allows the transferred electron greater delocalization Thus, the underestimation of coulomb repulsion would promote such charge transfer

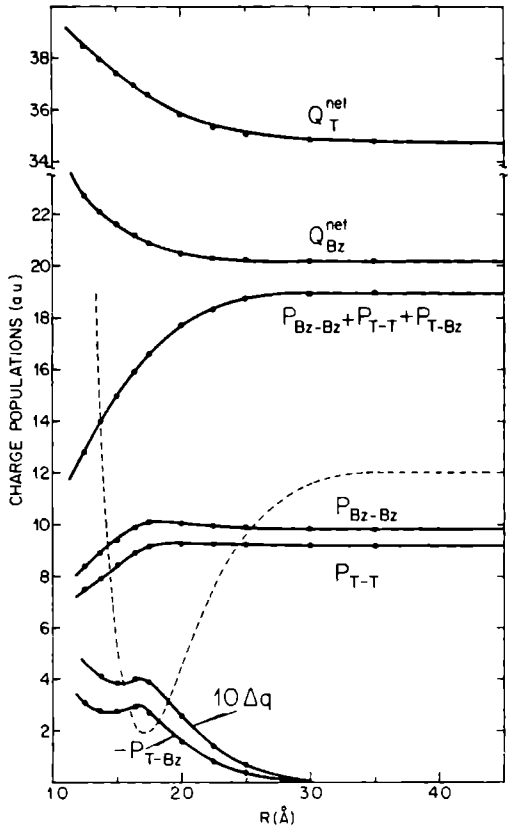
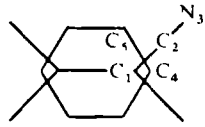


Fig 2 Various summed overlap and net charge populations and the amount of charge transfer (all in a u) as a function of the interplanar separation R for the non rotated structure See text for definition of symbols

Also plotted in Fig 1 is the dipole moment, μ, which was calculated directly in the CNDO approximation Here again the value at the calculated equilibrium separation (about 3 5 debyes) is larger than the generally expected moments (of the order of 1 to 1 5 debyes) The large dipole moments is attributed to the large degree of charge transfer found in the calculations The behavior of the dipole moment and also the charge populations shown in Fig 2 as one approaches the equilibrium separation and proceeds to shorter distances indicates clearly the onset of the repulsive interactions which have clearly changed the behavior of the electrons in the complex Although μ tends to drop ultimately, Δq, the amount of charge transfer, continues to increase as shorter distances are approached Were we to take our interpolated results at, say, 3 25 Å, we obtain at ΔE of $-3\,3$ kcal/mole and a dipole moment of about 0 1 Debye Clearly one must be cautious in handling results obtained from a single geometry

The search for possible intermolecular bonding was one of the initial reasons for initiating this study and was the reason for performing the Mulliken population

analyses. For $R \geqq 1.75$ Å *no* positive intermolecular overlap populations (OP's) are found; indeed, all values are in the range generally expected for non-bonded atoms.



At $R = 1.75$ Å the C_2–C_5 OP is -0.078, the largest magnitude observed, while the C_1–C_4 OP is -0.018 One might have expected this latter OP to be most indicative of bonding; perhaps the fact that it is the least negative of the significant OP's does hint at this fact. All of the atoms "near" each other exhibit similar negative OP's and this includes the hydrogen interactions (which would indeed normally be expected to show repulsive effects).

To indicate some general trends in the charge populations we have plotted certain sums for the non-rotated structure in Fig. 2. P_{T-T} and P_{Bz-Bz} represent the *sums* of OP's between atoms (all atoms) on TCNE and, separately, on benzene, respectively; P_{T-Bz} represents the summed OP's between all atoms on the two different molecules. Q_T^{net} and Q_{Bz}^{net} are the *net* Mulliken charges summed for the separate molecules and Δq represents the total charge transferred *from benzene to* TCNE; a value of $+1$ for Δq would represent the system $(TCNE)^{-1}$ $(benzene)^{+1}$. Δq is calculated by dividing P_{T-Bz} equally between the two molecules and represents the increase in the total Mulliken gross charge on TCNE.

The energy curve is indicated in the background for reference. There are several interesting aspects. The P_{T-T} and P_{Bz-Bz} sums hardly change (slight increase) as R decreases, indicating that the net intramolecular binding is not being much affected by those forces which are stabilizing the complex. The total OP $(P_{T-T} + P_{Bz-Bz} + P_{T-Bz})$ does show a uniform decrease throughout the range of decreasing R while both net charge sums increase We interpret this to indicate that the stabilization arises essentially solely from charge transfer. Nearly all the individual atomic net charges increase, except for the hydrogen atoms which show a decrease.

Decreasing overlap populations are indicative of the increase of anti-bonding MO's in the total wave function. Such anti-bonding orbitals will tend to pile charge up on atoms as net charges at the expense of bond populations. This is apparently what is happening in our present calculation. The intermolecular OP's are being driven negative causing the net atomic charges to increase on *both* molecules The distribution is such as to favor the more electron-accepting TCNE molecule giving rise to a finite Δq and dipole moment. The parallel nature of $-P_{T-Bz}$ and Δq would seem to indicate that it is basically only the intermolecular coupling which is anti-bonding (in the sense of negative OP's) and that the MO's of the whole complex are the MO's of the individual molecules weighted (polarized) toward the TCNE molecule. This fits in with the essentially constant behavior of P_{T-T} and P_{Bz-Bz} (down to $R = 1.75$ Å, at least) which indicates that molecular integrity is maintained for each molecular component. Although the calculated energy and equilibrium distance might indicate the formation of one "super molecule", the Mulliken population analysis shows clearly that this is not what

is happening; rather, the two component molecules remain easily recognizable as such, down to the shortest distance calculated (1.25 Å).

Obviously a careful study of the dominant parameters in the CNDO approach should be made in order to make a proper evaluation of the method as applied to these systems. If London forces are indeed the significant factor perhaps more than a simple SCF approach will be necessary. However, if one is to accept both the present method of calculation and the premise that SCF wavefunctions are "reasonably good" one must conclude that charge transfer may yet prove to be a significant factor in the stability of molecular complexes.

Acknowledgement. We are indebted to Duke University and the Duke University Computational Center for partial support of our computer calculations and for technical assistance.

Appendix

In this appendix we will show that equivalent bonds in a molecule are not necessarily characterized by identical bond orders but that summed overlap populations do reflect the physical equivalence of symmetry-related bonds. The latter summed, or total, overlap populations are thus better suited to characterize chemical bonds within a molecule.

In most computer work the localized atomic orbitals (AO's) are determined by the local coordinate system whereas at other times one often employs equivalent AO's (characterized by the fact that symmetry operations induce essentially permutations of the equivalent AO's among themselves). We consider initially an arbitrary set of AO's.

The charge and bond order matrix, $\boldsymbol{P}$, for a single determinantal wave function constructed from orthonormal molecular orbitals is invariant under transformations of the subspace spanned by the occupied MO's. A symmetry operation g mixes only MO's of equal energy; thus, for a closed-shell system, symmetry operators transform only occupied MO's among themselves and therefore leave $\boldsymbol{P}$ invariant.

$$g\boldsymbol{P} \equiv \boldsymbol{P}' = \boldsymbol{P} . \tag{1}$$

To describe the symmetry operation g with respect to the AO's let $|\alpha i\rangle$ be the i^{th} AO on center α; then in general

$$g\,|\alpha i\rangle = \sum_{\beta j} |\beta j\rangle\, D(g)_{\beta j;\,\alpha i} . \tag{2}$$

Since g is a one-to-one mapping of the set of nuclei onto itself, we can write

$$D(g)_{\beta j;\,\alpha i} = \delta_{\beta\alpha'} D^{(\alpha)}(g)_{ji} \qquad g : \alpha \to \alpha' \tag{3}$$

in which α' is the g-image of α. If the atomic orbitals span representations of the full rotation group, then the matrix $\boldsymbol{D}^{(\alpha)}(g)$ will represent the rotation of the local axes on α' relative to the g-image of the α-coordinate system. We will assume that this is the case, and furthermore that the representation is orthogonal.

Consider next how the filled MO's $\{|s\rangle\}$ transform under g, let $|s\rangle$ be the MO

$$|s\rangle = \sum_{\alpha i} C_{\alpha i\ s} |\alpha i\rangle$$

Then

$$\begin{aligned} g|s\rangle &= \sum_{\alpha i} C_{\alpha i\ s} g|\alpha i\rangle \\ &= \sum_{\alpha i}\sum_{\beta j} C_{\alpha i\ s} D(g)_{\beta j\ \alpha i} |\beta j\rangle \\ &= \sum_{\beta j} C'_{\beta i\ s} |\beta j\rangle \end{aligned} \tag{4}$$

Where the second and third lines of Eq (4) define the quantity $C'_{\beta j\ s}$ Thus, assuming for simplicity real coefficients,

$$\begin{aligned} P'_{\alpha i\ \beta j} &= \sum_{s} C'_{\alpha i\ s} C'_{\beta j, s} \\ &= \sum_{\gamma k}\sum_{\delta p}\sum_{s} D(g)_{\alpha i\ \gamma k} D(g)_{\beta j\ \delta p} C_{\gamma k\ s} C_{\delta p\ s} \end{aligned} \tag{5}$$

Where, from the definition of the set $\{|s\rangle\}$, the sum over s implies a sum over the occupied MO's of the system
The matrix representation $\boldsymbol{D}(g)$ of g is orthogonal, that is

$$D(g)_{\alpha i\ \gamma k} = D(g^{-1})_{\gamma k\ \alpha i} \tag{6}$$

Put $h = g^{-1}$ and define α'' by $h\ \ \alpha \rightarrow \alpha''$

$$D(g)_{\alpha i\ \gamma k} = D(h)_{\gamma k\ \alpha i} = \delta_{\gamma\alpha}\ D^{(\alpha)}(h)_{ki}\,. \tag{7}$$

Eq (1) indicates that $\boldsymbol{P} = g\boldsymbol{P} = \boldsymbol{P}'$, so

$$\begin{aligned} P_{\alpha i\ \beta j} &= \sum_{\gamma k}\sum_{\delta p} \delta_{\gamma\alpha}\ D^{(\alpha)}(h)_{ki}\delta_{\delta\beta}\ D^{(\beta)}(h)_{pj} P_{\gamma k\ \delta p} \\ &= \sum_{k\ p} D^{(\alpha)}(h)_{ki} D^{(\beta)}(h)_{pj} P_{\alpha\ k\ \beta\ p} \end{aligned} \tag{8}$$

We conclude that $P_{\alpha i\ \beta j}$ is in general not equal to $P_{\alpha\ i\ \beta\ j}$ in which α' and β'' denote atoms equivalent to, respectively, α and β If we sum over i and j in order to obtain the total bond order, we will get different answer for the bond orders of the α-β and the α''-β'' bond, [unless, of course, $\boldsymbol{D}^{(\alpha)}(h)$ and $\boldsymbol{D}^{(\beta)}(h)$ are identity matrices, meaning that the local coordinate systems transform into one another under h]

We now prove the validity of the following relation between summed Mulliken overlap populations

$$\sum_{i\ \text{on}\ \alpha}\ \sum_{j\ \text{on}\ \beta} P_{\alpha i\ \beta j} S_{\alpha i\ \beta j} = \sum_{i\ \text{on}\ \alpha}\ \sum_{i\ \text{on}\ \beta} P_{\alpha\ i\ \beta\ j} S_{\alpha\ i\ \beta\ j} \tag{9}$$

First we transform S with h, since S is invariant under a unitary transformation:

$$\begin{aligned} S_{\beta_j,\alpha_i} = S''_{\beta_j,\alpha_i} &= \sum_{\gamma k}\sum_{\delta p} D(h)_{\gamma k,\beta_j} D(h)_{\delta p,\alpha_i} S_{\gamma k,\delta p} \\ &= \sum_{k,p} D^{(\beta)}(h)_{kj} D^{(\alpha)}(h)_{pi} S_{\beta' k,\alpha'' p} \\ &= \sum_{k,p} D^{(\beta)}(h)^{-1}_{jk} D^{(\alpha)}(h)^{-1}_{ip} S_{\beta'' k,\alpha'' p} . \end{aligned} \tag{10}$$

Substituting the results (8) and (10) in the l. h.s. of (9) and using the fact that S is symmetric we obtain:

$$\begin{aligned} \sum_{i,j} P_{\alpha_i,\beta_j} S_{\alpha_i,\beta_j} &= \sum_{i,j}\sum_{k,p}\sum_{q,r} D^{(\alpha)}(h)_{ki} D^{(\alpha)}(h)^{-1}_{ir} D^{(\beta)}(h)_{pj} D^{(\beta)}(h)^{-1}_{jq} P_{\alpha'' k,\beta'' p} S_{\alpha'' r,\beta'' q} \\ &= \sum_{k,p} P_{\alpha\ k,\beta\ p} S_{\alpha'' k,\beta'\ p} . \end{aligned} \tag{11}$$

This result indicates that the overlap population between two atoms summed over the atomic orbitals of the atoms is identical for equivalent pairs of atoms.

If one is dealing with equivalent atomic orbitals, the matrix $D^{(\alpha)}(g)$ takes a simple form. Suppose g transforms $|\alpha i\rangle$ into $E_{ii}\,|\alpha' i'\rangle$, in which $E_{ii'} = \pm 1$ (if we use real AO's). $D^{(\alpha)}(q)$ takes the form:

$$D^{(\alpha)}(g)_{ij} = \delta_{ji'} E_{ii} . \tag{12}$$

Substituting this in (8):

$$P_{\alpha_i,\beta_j} = E_{ii}\, E_{jj'} P_{\alpha' i',\beta' j'} \tag{13}$$

and in (10):

$$S_{\alpha_i,\beta_j} = E_{ii}\, E_{jj'} S_{\alpha' i',\beta' j'} . \tag{14}$$

From this we see that symmetry related elements of P and S differ at most in sign; further the product equality

$$P_{\alpha_i,\beta_j} S_{\alpha_i,\beta_j} = P_{\alpha i',\beta' j'} S_{\alpha i',\beta' j'} \tag{15}$$

will hold.

References

1 Dewar, M J S, Thompson, C C, Jr Tetrahedron Suppl No 7, 97 (1966)
2 Chesnut, D B, Moseley, R W Theoret chim Acta (Berl) **13**, 230 (1969)
3 Pople, J A, Beveridge, D I Approximate molecular orbital theory New York McGraw-Hill 1970, and references therein
4 Momany, F A, McGuire, R F, Yan, Y F, Scheraga, H A J physic Chem **74**, 2424 (1970).
5 Kollman, P A, Allen, L C J Amer chem Soc **92**, 753 (1970)
6 Prout, C K, Wright, J D Angew Chem (Int Ed) **7**, 659 (1968)
7 Williams, R M, Wallwork, S C Acta crystallogr **22**, 899 (1967)
8 Bekoe, D A, Trueblood, K N Z Kristallogr **113**, 1 (1960)
9 Mulliken, R S J chem Physics **23**, 1833, 1841 (1955)
10 Coulson, C A Proc Roy Soc (London) A **169**, 413 (1939)
11 Lowdin, P O J chem Physics **18**, 365 (1950)
12 Carlson, B C, Keller, J M Physic Rev **105**, 102 (1957)
13 Brown, R D, Roby, K R Theoret chim Acta (Berl) **16**, 175 (1970)

14 Fraga, S, Ransil, B J J chem Physics **34**, 727 (1961)
15 a) Newton, M D, Boer, F P, Lipscomb, W N J Amer chem Soc **88**, 2368 (1966)
b) Palke, W E, Lipscomb, W N J Amer chem Soc **88**, 2384 (1966)
16 Briegleb, G, Czekalla, J, Reuss, G Z physik Chem (Neue Folge) **30**, 334 (1961)
17 Kroll, M J Amer chem Soc **90**, 1097 (1968)
18 Lippert, J L, Hanna, M W, Trotter, P J J Amer chem Soc **91**, 4035 (1969)
19 Murrell, J N, Randic, M, Williams, D R Proc Roy Soc (London) A **284**, 566 (1965)
20 Herndon, W C, Feuer, J J Amer chem Soc **90**, 5914 (1968)
21 Mantione, M J Theoret chim Acta (Berl) **15**, 141 (1968)
22 Jesaitis, R G, Streitwieser, A Theoret chim Acta (Berl) **17**, 165 (1970)

Dr D B Chesnut
Department of Chemistry
Duke University
Durham, North Carolina 27700, USA

Ab-initio valence bond calculations on the He–He potential curve using small bases†

by PAUL E. S. WORMER, TON VAN BERKEL and AD VAN DER AVOIRD
Institute of Theoretical Chemistry, University of Nijmegen
Nijmegen, The Netherlands

(*Received* 20 *May* 1974)

In studying methods for the *ab-initio* calculation of Van der Waals interactions which can be extended to larger molecules, we have tested the Valence Bond scheme on the He_2 system. This method looks promising, as it appears to yield in one consistent formalism both the attractive dipole–dipole terms and the repulsive exchange terms with reasonable accuracy, using only a simple orbital basis and a very small number of VB structures.

1 Introduction

The interaction between two ^1S ground-state He atoms has been investigated by many people with methods of varying degrees of accuracy and sophistication. The older approaches (e.g. reference [1]) use perturbation formalisms which account for the long-range Van der Waals attraction, and superpose empirical repulsive potentials in an *ad hoc* manner, in order to give good fits to the depth and position of the Van der Waals minimum. Later workers [2–4] concentrated on the exchange interaction, mainly using an MO scheme. These investigators were not able to obtain minima in the potential curves, so one started to realize that the computation within a uniform model of a complete interaction curve for He_2 poses a real problem. Even the 64-term Configuration Interaction wave function of Phillipson [5] did not predict the correct depth of the Van der Waals minimum. One still had to obtain the repulsive and attractive contributions to the energy by separate calculations [6]. This situation remained unaltered all through the sixties [7–10], which led to the now common conviction that a single configuration potential curve for two closed-shell atoms must be entirely repulsive [11].

In 1970 two independent simultaneous publications [12, 13] presented high quality potential curves, each calculated by a single consistent computational technique. Although these results prove beyond doubt that the Van der Waals potential well can be computed with an accuracy comparable to that of experiment, it will be difficult to apply these methods to larger systems.

The purpose of this communication is to point out that reasonable results can be obtained by very simple calculations, that is, simple with regard to the size of the basis sets and the computer time spent on the whole curve. Only if one is

† The investigations were supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

able to predict with minimal effort the Van der Waals energy of a small system like He–He will it be possible to hope for positive results in computations on larger systems.

2. Method

The Valence Bond (VB) method is similar to a perturbation treatment in the sense that it is based on the states of the unperturbed atoms. In VB theory, one solves a secular problem in which an element of the hamiltonian matrix has the general form

$$H_{i,j} = \langle \mathscr{A} P_S \Psi_{i'}^{A} \Psi_{i''}^{B} H \mathscr{A} P_S \Psi_{j'}^{A} \Psi_{j''}^{B} \rangle. \tag{1}$$

Here $\mathscr{A}$ is the antisymmetrizer for the total system, P_S a spin-adapting projection operator and $\Psi_{i'}^{A}$, $\Psi_{i''}^{B}$ are unperturbed states of the atoms A and B respectively.

The two main problems in the computation of $H_{i,j}$ are the effect of P_S and the non-orthogonality between the orbitals on the atoms A and B. The matrix elements (1) are constructed by Reeves' algorithm [14] which is based on a Young projector [15].

This algorithm expects orthonormal spatial orbitals, and so we orthonormalize a chosen set of atomic orbitals by means of a matrix $\mathbf{t}$. Then the full H-matrix $\mathbf{H}$ with elements defined by (1) is constructed on the basis of *all* states (with a certain spin multiplicity) which arise from the orthonormalized AO set. The effect of the orbital transformation $\mathbf{t}^{-1}$ on the N-electron wave functions is given by an Nth-order tensor representation [16] $\mathbf{T}(\mathbf{t}^{-1})$ which undoes the effect of $\mathbf{t}$ after performance of the transformations

$$\mathbf{O} = \mathbf{H}'\mathbf{C}' - \mathbf{S}'\mathbf{C}'\mathbf{E} = \mathbf{T}^{\dagger}(\mathbf{t}^{-1})\mathbf{H}\mathbf{T}(\mathbf{t}^{-1})\mathbf{T}(\mathbf{t})\mathbf{C} - \mathbf{T}^{\dagger}(\mathbf{t}^{-1})\mathbf{S}\mathbf{T}(\mathbf{t}^{-1})\mathbf{T}(\mathbf{t})\mathbf{C}\mathbf{E}. \tag{2}$$

Specific VB structures can be selected by considering only certain columns of $\mathbf{T}(\mathbf{t}^{-1})$ and so the dimension of the primed secular equation, which is on the basis of the original non-orthogonal AO's, is generally smaller than the dimension of the H-matrix defined by (1). Solution of the transformed equation in the usual manner gives the desired VB states and energies.

In this procedure, which we intend to elaborate in a future publication, one avoids tedious derivations of matrix-elements on basis of non-orthogonal orbitals§.

The exact form chosen for the matrix $\mathbf{t}$ is unimportant if the effect of the orthonormalization is annihilated according to equation (2). However, for large problems this is not feasible and in that case the choice of a good orthonormalization procedure [17] becomes important. We have promising experiences with the following three-step process:

(i) Löwdin-orthonormalize the orbitals occupied in the free atoms.
(ii) Perform a Gram Schmidt orthonormalization of the remaining orbitals onto the occupied ones.
(iii) Lowdin-orthonormalize the orbitals obtained in step (ii) among each other.

§ We thank the referee for drawing our attention to two papers [25, 26] which treat the non-orthogonality problem in a similar way. In these references the secular matrices are defined on the basis of Slater determinants instead of Young projected functions. Although the use of Slater determinants facilitates the construction of $\mathbf{T}(\mathbf{t}^{-1})$, it requires a much larger dimension of the secular problem, since one does not work with eigenfunctions of S^2.

Programmes based on the outlined ideas have been written for an IBM 370/158 computer, taking as integral input the output from the IBMOL–5 programme [18].

3. Results

On each He atom the four orbitals: 1s, $2p_x$, $2p_y$, $2p_z$ are placed. The 1s orbitals (SCF orbitals from Huzinaga [19]) are represented by a contracted set of six primitive gaussians. Also six primitives are contracted to represent a single-ζ Slater-type p orbital. The exponent ζ_p of this STO can be optimized in four different ways, see table 1.

Energy minimized	ζ
Correlated atomic ^1S energy	2·45
Excited state SCF ^3P energy	0·58
Excited state SCF ^1P energy	0 47
Van der Waals energy He He at $R = 5{\cdot}6$ bohr	1 3

Table 1. Optimized STO exponents ζ_p of the 2p orbital.

The value $\zeta_p = 1{\cdot}3$ compares well with the value $\zeta_p = 1{\cdot}378$ found from an analytic optimization of the Van der Waals attraction [20]; the other values in table 1 agree with those given in reference [12]. In all subsequent calculations the value of ζ_p is held fixed on 1·3. The first calculations are based on the four σ-type AO's only. Although the usual perturbation approach [11] starts from purely co-valent states, yet we started with a set of 'full' VB calculations. Since four AO's give rise to at most 20 singlet structures, full VB implies a 20-dimensional secular problem. The results are shown in figure 1.

Compared to the experimental values [21]: $E_m/k = 11{\cdot}0$ K and $R_m = 5{\cdot}6$ bohr† the minimum is much too deep. Inspection of the wave function Ψ at the minimum, while retaining only the dominant terms

$$\Psi = 0{\cdot}999\,[s_A^2, s_B^2] - 0{\cdot}034\,\{[s_A^2, p_B^2] + [p_A^2, s_B^2]\} - 0{\cdot}005\,\{[s_A^2, p_A, p_B] + [p_A, p_B, s_B^2]\} + 0{\cdot}004\,[s_A, p_A, s_B, p_B]$$

shows that the leading structure represents the Hartree-Fock uncorrelated ground state in the separated atom limit. The last term in Ψ is the one predicted by perturbation theory to be responsible for the induced dipole-induced dipole part of the dispersion energy [11]. The second structure $[s_A^2, p_B^2]$ is known to contribute considerably to the intra-atomic correlation energy of atom B, and $[p_A^2, s_B^2]$ does the same for atom A. But we expect a more or less constant shift of the potential curve from these latter terms, as the intra-atomic correlation energy does not depend heavily on R [22]. This is confirmed by the coefficients of the first three structures which are practically independent of R up to the separated atom limit.

Unexpected is the presence of the excited CT (charge transfer) structures $[s_A^2, p_A, p_B]$ and $[p_A, p_B, s_B^2]$ while the CT states $[s_A^2, p_A, s_B]$ and $[s_A, p_A, s_B^2]$ are

† $k = 1{\cdot}38054 \times 10^{-23}$ J K (Boltzmann's constant). 1 bohr $= 5\,29167 \times 10^{-11}$ m.

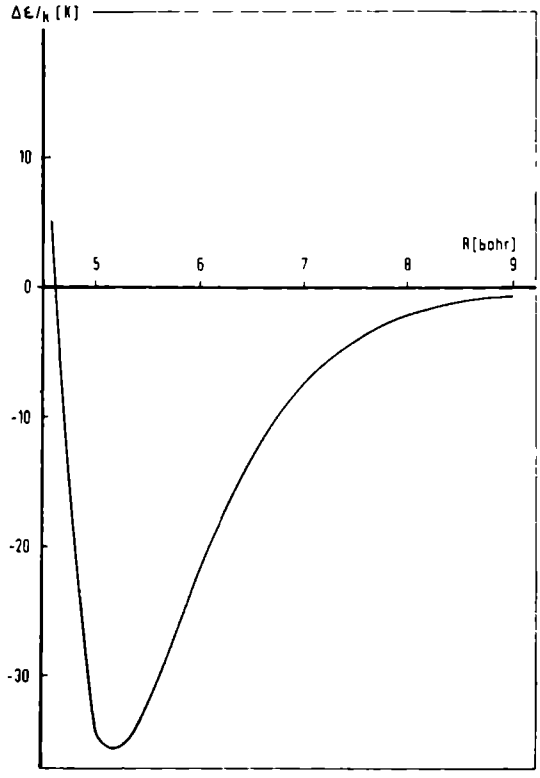


Figure 1. Interaction energy in He_2 calculated by full VB (20 structures) on four σ orbitals.

absent. After a moment of thought the relative importance of these terms can be explained as follows. Making the formal expansion of $p_A(\mathbf{r}_A)$ in a set of atomic functions centred on B:

$$p_A(\mathbf{r}_A) = \sum_{n,l} C_{nl}(\mathbf{R}) f_{nl}(\mathbf{r}_B)$$

we write

$$[s_A^2, p_A, s_B] = \sum_{n,l} C_{nl}(\mathbf{R})[s_A^2, f_{nl}(\mathbf{r}_B), s_B(\mathbf{r}_B)].$$

The right-hand side of this equation is an expansion in singly excited states on atom B, which, $s_B(\mathbf{r}_B)$ being an SCF orbital, does not contribute to the ground-state via intra-atomic terms in the hamiltonian operator. (Even in the case of non-orthogonal orbitals [27] it can be proved easily, by invoking Brillouin's theorem, that the admixture of singly excited states leaves the ground state of atom B a pure $[s_B^2]$ state.) This explains the unimportance of CT terms without simultaneous excitation.

In the same way the excited CT structure becomes

$$[s_A^2, p_A, p_B] = \sum_{n,l} C_{nl}(\mathbf{R})[s_A^2, f_{nl}(\mathbf{r}_B), p_B(\mathbf{r}_B)],$$

from which we deduce immediately that this structure yields a contribution to the intra-atomic correlation energy of atom B Because of the R-dependent expansion coefficient C_{nl} ($\mathbf{R}$) this contribution is distance dependent To get a better grasp of this distance-dependent improvement of the intra-atomic correlation energy we performed VB calculations on the free He atom A, with a basis of structures on the 'empty' (no electrons, no nucleus) centre B approaching A The results are plotted in figure 2 The influence of the basis on B is seen to grow drastically with decreasing distance It is important to note that the He atom is *physically* unperturbed, and that this artificial energy lowering is a *mathematical* effect due to the enlargement of the structure basis Around 5 5

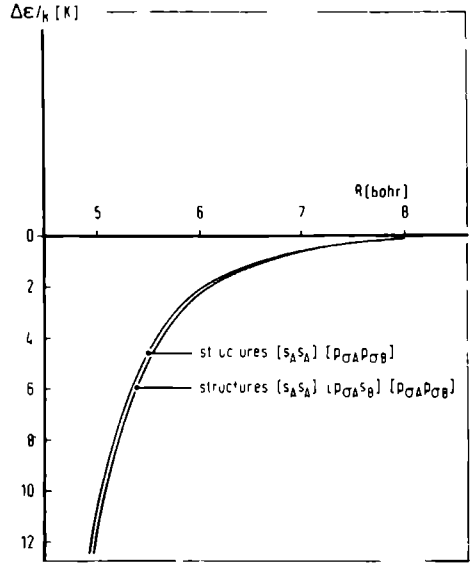


Figure 2 Distance dependency of the energy of the He atom A in the structure basis of A+B multiplied by 2 Only the structures with non-negligible contribution are indicated

bohr this energy lowering is of the same order of magnitude as the physical interaction between two He atoms, but together with the physical interaction this energy does not fully add up to the well depth of the full VB calculation So we must conclude that the effect of the basis set enlargement is not an additive one and, therefore, to eliminate this phenomenon we omit all structures that contribute to the intra-atomic correlation, in accordance with the purpose of this investigation† In figure 3 three computed potential curves are exhibited The

† After completion of this work a paper on the He–He interaction [28] appeared which also considers in detail the effect of basis enlargement The authors use a larger atomic orbital and configuration basis They perform a large Configuration Interaction calculation and correct for the distance-dependent basis-set effect by subtracting the result of an atomic calculation with a vacant basis on the other centre, thus assuming additivity

upper two are calculated with σ-type AO's, taking into account the two co-valent structures $[s_A^2, s_B^2]$ and $[s_A, p_A, s_B, p_B]$ only. The lower of these two is obtained from pure, non-orthogonal AO's, and the higher one from AO's that are orthogonalized by the three-step process described in the previous section. Notice that because of the careful orthogonalization the curves almost coincide.

We have tried several other ways of orthonormalization, for instance SCF followed by localization, but they all give results that are definitely inferior to the ones shown.

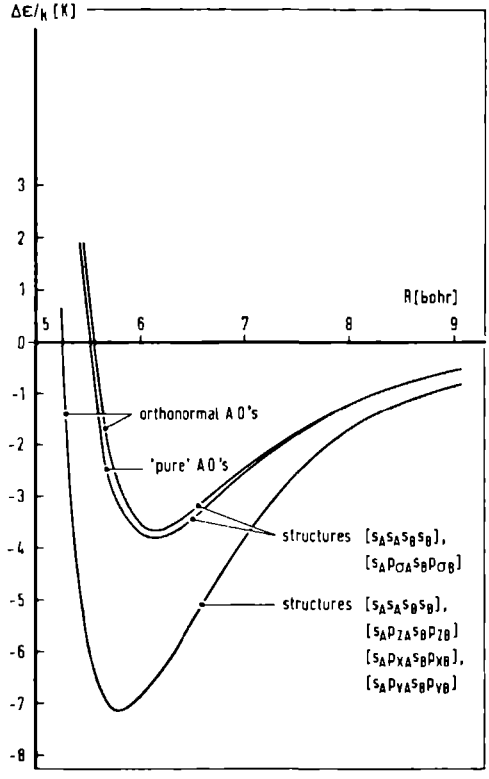


Figure 3. Van der Waals energy calculated by VB on the basis of four σ orbitals (two structures) or eight σ + π orbitals (four structures).

The lowest of the three curves in figure 3 has been computed by inclusion of two π AO's on each atom and the four structures:

$$[s_A^2, s_B^2],\ [s_A, p_{zA}, s_B, p_{zB}],\ [s_A, p_{yA}, s_B, p_{yB}],\ [s_A, p_{xA}, s_B, p_{xB}].$$

Here the AO's are orthogonalized, again following the described three-step scheme.

All doubly-excited configurations give rise to two linearly independent singlet VB structures. One of these structures corresponds to the coupling of

two excited singlet He atoms, the other one to the coupling of two excited triplet atoms [15] It appears that the contributions of the latter structures to the interaction energy are negligible

4 Discussion

In figure 3 is shown how a very simple calculation based on two σ-type AO's per atom and only two VB structures is capable of giving already an important part of the experimental Van der Waals energy (upper two curves) One cannot expect much more, of course, since in effect only the σ-dependent part of the dipole–dipole interaction [11] is taken into account

Including the π contributions to the dispersion energy by adding just two more structures it can be concluded that a very simple computation suffices to give a major part of the Van der Waals energy Here the theoretical limit is the total induced dipole-induced dipole (R^{-6}) contribution to the dispersion energy Kestner [11] estimated this to give a minimum of 5 68 K, but this agrees neither with the results of Schaefer *et al* [12, 23], nor with ours

In table 2 our results are compared with those of Schaefer *et al* , the agreement, especially in the outer region, is striking, since these authors based their $\sigma+\pi$ calculations on five s and pσ orbitals and four pπ orbitals per atom, yielding 72 σ configurations and 16 π configurations

	SCF	$\langle s_A^2 s_B^2 \vert H \vert s_A^2 s_B^2 \rangle$	σ(s+p)		$\sigma+\pi$(s+p)	
R bohr	Ref [23]	This work	Ref [23]	This work	Ref [23]	This work
5 2	24 45	23 43	+7 7	+8 08	−0 6	+0 57
5 4	15 03	14 11	+1 7	+1 64	−5 0	
5 6	9 24	8 46	−1 6	−1 71	−7 0	−6 63
5 8	5 66	5 05	−3 08	−3 25	−7 46	−7 26
6 0	3 45	3 00	−3 64	−3 82	−7 17	−7 14
6 2		1 77		−3 85		−6 60
6 4		1 04		−3 60		−5 90
6 5	1 00		−3 38		−5 56	
6 6		0 63		−3 25		
6 8		0 32		−2 87		−4 48
7 0	0 287	0 22	−2 51	−2 49	−3 90	
8 0	0 023	0 03	−1 23	−1 17	−1 85	−1 83
9 0	0 002	0 00	−0 61	−0 57	−0 92	
10 0	0 000	0 00	−0 33	−0 28	−0 49	
12 0		0 00		−0 09		
14 0		0 00		0 00		−0 06

Table 2 Interaction energies in units of degrees Kelvin The SCF energy for two isolated He atoms is −5 722 232 8 hartree

Concluding, it can be stated that the VB method performs very well for weakly interacting systems such as two He atoms It constitutes a single consistent model capable of giving a good prediction of the entire Van der Waals well.

Of course, in this work the intra-atomic correlation correction is neglected by using a Hartree-Fock atom as the reference system At the Van der Waals minimum, however, this correction differs only 1·23 K from the correlation energy

in the free atoms [23] Although this difference increases for shorter interatomic distances [24], it remains rather small relative to the repulsion energy Moreover, the VB scheme can only be expected to perform well if the atoms do not deviate too much from their electronic structure at infinite distance If one is interested in the repulsive region at smaller distances, an approach based on the Hartree Fock molecule as the reference system seems more appropriate, unless one is prepared to include substantially larger numbers of VB structures than we did in this paper

The authors thank Drs E Clementi and F. B. van Duijneveldt for making available the IBMOL-5 programme and Dr M van Hemert for the four-index transformation programme One of us (Wormer) thanks Dr C Moser for his hospitality at C E C A M , where part of the programme was written

References

[1] Margenau, H , 1931, *Phys Rev* , **38,** 747
[2] Griffing, V and Wehner, J F , 1955, *J. chem Phys* , **23,** 1024
[3] Sakamoto, M , and Ishiguro, E , 1956, *Prog theor Phys* , **15,** 37
[4] Huzinaga, S , 1958, *Prog theor Phys* , **20,** 15
[5] Phillipson, P E 1962, *Phys Rev* , **125,** 1981
[6] Murrell, J N , and Shaw, G , 1967, *Molec Phys* , **12,** 475
[7] Ransil, B J , 1961, *J chem Phys* , **34,** 2109
[8] Kestner, N R , 1968, *J chem Phys* , **48,** 252
[9] Gilbert, T L , and Wahl, A C , 1967, *J chem Phys* , **47,** 3425
[10] Schwartz, M E , and Schaad, L J 1968, *J chem Phys* , **48,** 4709
[11] Margenau, H , and Kestner, N R , 1969, *Theory of Intermolecular Forces* (Pergamon)
[12] Schaffer III H F , McLaughlin, D R , Harris, F E , and Alder, B J , 1970, *Phys Rev Lett* , **25,** 988
[13] Bertoncini P , and Wahl A C , 1970, *Phys Rev Lett* , **25,** 991
[14] Reeves C M , 1966, *Commun A C M* **9,** 276
[15] Wormer, P E S , and van der Avoird A , 1974 *Int J quant Chem* , **8,** 715
[16] Wormer P E S , and van der Avoird A , 1972, *J chem Phys* **57,** 2498
[17] Löwdin, P -O , 1970, *Adv quant Chem* , **5,** 185
[18] Clementi, E , and Mehl, J , 1971, *IBM Rep* RJ 883
[19] Huzinaga, S , 1965, *J chem Phys* , **42,** 1293
[20] Kockel, B , and Wirsam B , 1971, *Z Naturf* A, **26,** 107
[21] Farrar, J M , and Lee, Y T , 1972, *J chem Phys* , **56,** 5801
[22] Bertoncini, P J , and Wahl, A C , 1973, *J chem Phys* , **58,** 1259
[23] McLaughlin, D R , and Schaefer, III, H F , 1971, *Chem Phys Lett* , **12,** 244
[24] Van der Velde, G A , 1974, Thesis, Groningen, The Netherlands
[25] Moffitt, W , 1953, *Proc R Soc* A, **218,** 486
[26] Balint-Kurti, G G , and Karplus, M , 1969, *J chem Phys* , **50,** 478
[27] Grein, F , and Chang, T C , 1971, *Chem Phys Lett* , **12,** 44
[28] Daudey, J P , Malrieu, J P , and Rojas, O , 1974, *Int J quant Chem* , **8,** 17

Ab initio valence-bond calculations of the van der Waals interactions between π systems: The ethylene dimer*

Paul E S Wormer and Ad van der Avoird

Institute of Theoretical Chemistry University of Nijmegen The Netherlands
(Received 3 December 1974)

A multistructure valence-bond method for the calculation of van der Waals forces is presented which includes in one consistent formalism the electrostatic, induction, and dispersion forces and takes exchange correctly into account The application of this method to the ethylene dimer leads to the following main conclusions (1) The "first order" electrostatic forces are comparable in magnitude to the "second order" forces even though the molecules possess no permanent dipole moments Dispersion forces are much larger than induction Second order interactions are more isotropic than first order forces (2) In the multipole expansions of the long range forces, the inclusion of the first term only is not sufficient for a good approximation to the interaction (3) Exchange effects become nonnegligible at approximately 12 bohr, while the van der Waals minimum between two perpendicular molecules is at 9 4 bohr At about 6 bohr, penetration effects make the multipole expansion meaningless Possible simplifications for future calculations are suggested A seemingly good van der Waals minimum which is calculated by the *ab initio* SCF method is shown to be caused in part by the (mathematical) effect of basis set enlargement CNDO results disagree completely with the *ab initio* calculations

I. INTRODUCTION

Van der Waals forces between π-electron systems are of great importance in accounting for the conformation of many of the biopolymers.[1] Together with hydrogen bonding, these forces play a decisive role in the secondary structure of the nucleic acids and proteins. Many authors[2-5] have discussed the interaction between the π electrons in the stacked bases of the double stranded DNA molecule.

Similar stacks of molecules are found in all solid CT (charge transfer) complexes of the $\pi-\pi$ type.[6] In CT complexes,[7-11] as well as in the majority of the molecular crystals,[12-15] van der Waals forces can be held responsible for the stabilization energy.

Most of the calculations on the long and intermediate range interactions between organic molecules are based on the ideas of London,[16,17] who applied Rayleigh–Schrödinger first and second order perturbation theory. In his earlier work,[16] London proposed a multipole expansion of the interaction operator, whereas later[17] he suggested rather to approximate the charge distributions on the monomers Both London schemes neglect intermolecular exchange and hence lack completely the Pauli exchange interactions. These are repulsive for closed-shell monomers and very dominant at shorter distances. Since the long range London forces are mostly attractive, inclusion of the Pauli repulsion is necessary in order to account for a van der Waals minimum. Several ways of dealing with the exchange have been proposed. On the one hand, perturbation formalisms have been derived which take exchange into account in all orders of perturbation,[18-20] but in general these methods are not practical for larger systems unless extra approximations are introduced. On the other hand, simpler and less rigorous ways of dealing with the exchange have been discussed,[13,21] all of which include parameters that are to be fitted to experiment.[22,23]

Although the semiempirical methods of computing intermolecular forces contain questionable approximations and assumptions,[23] they have been extremely useful in deepening the insight in the structure of biopolymers and molecular crystals; moreover, once the ambiguities in the approximations are removed, they are likely to remain the most fruitful approach in this field, since the more rigorous *ab initio* methods will in all probability stay too expensive to be applied to large systems in the foreseeable future.

Still, *ab initio* methods can be of great help in clearing up unsettled questions, as for instance the convergence of the multipole expansion,[24-26] the correctness of the segment monopole approximation,[27] or the nature of the exchange repulsion.[22]

Unlike the situation a decade or so ago, the approximations proposed by London[16,17] for computing the matrix elements occurring in a description of the van der Waals forces are no longer a necessity. Since the advent of high speed computers and sophisticated program packages (e.g., Ref. 28), one is not only able to produce good quality wavefunctions for the monomers, but also to calculate exactly all the intermolecular Coulomb and exchange integrals. Of course, this does not solve some of the more formal problems[18,19]: how to account for the exchange and how to approximate the infinite sum in the second order perturbation.

The first efforts made into the direction of *ab initio* calculations for intermolecular forces stayed within the SCF framework.[29,30] However, it soon became clear that the Hartree–Fock method does not yield dispersion energies,[31] but only electrostatic, induction, and exchange energies. So, a realistic calculation must include at least intermolecular correlation energy.[32,33]

In earlier work,[34] we have shown that it is feasible to calculate in the consistent formalism of the multistructure valence-bond method both the attractive dispersion and the repulsive exchange contribution to the total interaction energy. The valence-bond method is an ap-

pealing formalism in this respect, as it converges into London's theory for increasing intermolecular distances. Furthermore, a judicious choice of VB structures yields a quantitative description of the intermolecular correlation.

We have chosen to study the ethylene dimer because it is the simplest organic $\pi-\pi$ complex. As such, and also in its own right it has received much attention.[35-44]

II. THEORY

A. The valence-bond method and its connection with perturbation theory

In valence-bond theory, one calculates the energy of a system by solving a secular problem over "VB structures," which are antisymmetric many-electron eigenfunctions of the spin operator S^2. Spin-free VB structures may conveniently be constructed by acting with a Young operator Y on simple orbital products.[45] In this paper, where we are concerned with the interaction between two molecules A and B, the dimer VB structures are denoted by $Y(\psi_i^A \psi_j^B)$. Here ψ_i^A and ψ_j^B are monomer structures obtained from products of molecular orbitals localized on A and B, respectively, by projection with monomer Young operators. The monomer MO's used in this work are obtained from LCAO–SCF calculations on the separate monomers. It can be proved[46] that the dimer VB structures may also be constructed by acting directly with Y on simple products consisting of the same MO's as contained in ψ_i^A and ψ_j^B. In practice, we apply this latter procedure.

The monomer structures included in our calculations are the ground states ψ_0^A and ψ_0^B and singly excited states ψ_i^A and ψ_j^B. We expect the dimer structure $Y(\psi_0^A \psi_0^B)$ to account for the interactions between the unpolarized molecules, the structures $Y(\psi_i^A \psi_0^B)$ and $Y(\psi_0^A \psi_j^B)$ for the mutual induction effects and the doubly excited structures $Y(\psi_i^A \psi_j^B)$ for dispersion forces. This is clarified by the following discussion of the long range asymptotic behaviour of the valence-bond results.

It may be shown[46] that the contributions of the *intermolecular* permutations in Y to the Hamiltonian and overlap matrix elements start vanishing if the intermolecular distance R is increased. The VB structures $Y(\psi_i^A \psi_j^B)$ then become effectively equal to the products $\psi_i^A \psi_j^B$. These products form an orthonormal basis and the VB total energy for large R can be obtained by diagonalization of the total Hamiltonian on this basis The relation with Rayleigh–Schrödinger perturbation theory is shown if we write

$$H = H^A + H^B + V^{AB} \tag{1}$$

and imagine the diagonalization of H to be performed in two steps. First, diagonalize H^A in the basis $\{\psi_i^A\}$ by a linear transformation to $\{\phi_i^A\}$ so that

$$\langle \phi_i^A | H^A | \phi_{i'}^A \rangle = \delta_{ii'} E_i^A , \quad i = 0, 1, 2, \ldots \tag{2}$$

and analogously on monomer B:

$$\langle \phi_j^B | H^B | \phi_{j'}^B \rangle = \delta_{jj'} E_j^B , \quad j = 0, 1, 2, \ldots \tag{3}$$

Because of Brillouin's theorem, also valid in a finite LCAO model,[47] the SCF ground state ψ_0^A does not interact with the singly excited structures ψ_i^A, $i \neq 0$ under H^A, and therefore

$$\phi_0^A = \psi_0^A ,$$

and the same on B:

$$\phi_0^B = \psi_0^B .$$

After this first diagonalization step, we have the matrix elements

$$\langle \phi_i^A \phi_j^B | H | \phi_{i'}^A \phi_{j'}^B \rangle = \delta_{ii'} \delta_{jj'} (E_i^A + E_j^B) + \langle \phi_i^A \phi_j^B | V^{AB} | \phi_{i'}^A \phi_{j'}^B \rangle \tag{4}$$

The second step, which actually calculates the VB energy by completing the diagonalization of H, has to annihilate all the nondiagonal elements over V^{AB}. Assuming that the matrix elements over this operator are small, we can use the following expansion[48] for the lowest eigenvalue of H:

$$E_{VB} = E_0^A + E_0^B + \langle \phi_0^A \phi_0^B | V^{AB} | \phi_0^A \phi_0^B \rangle + \sum_{i,j}{}' \frac{|\langle \phi_0^A \phi_0^B | V^{AB} | \phi_i^A \phi_j^B \rangle|^2}{E_0^A + E_0^B - E_i^A - E_j^B} + \text{higher order terms.} \tag{5}$$

Because $\phi_0^A \phi_0^B$ is identical to the ground state $\psi_0^A \psi_0^B$, and the excited states can be regarded as the eigenstates of $H^{(0)} = H^A + H^B$ in the finite structure basis $\{\psi_i^A \psi_j^B\}$ this formula shows an explicit relation between perturbation theory and the valence-bond energy, valid for large intermolecular distances. We define for all distances

$$E_{VB}^{(1)} = \langle Y \psi_0^A \psi_0^B | H \; Y \psi_0^A \psi_0^B \rangle , \tag{6}$$

the expectation value of the total Hamiltonian over the dimer ground state VB structure, and

$$E_{VB}^{(2)} = E_{VB} - E_{VB}^{(1)} , \tag{7}$$

the energy lowering due to the inclusion of excited structures, and obtain expressions for the interaction energies,

$$\Delta E_{VB}^{(1)} = E_{VB}^{(1)} - E_0^A - E_0^B , \tag{8}$$

$$\Delta E_{VB}^{(2)} = E_{VB}^{(2)} , \tag{9}$$

which go asymptotically over into the first and second order perturbation energies. (The assumption that higher-order terms are small is confirmed by our calculations.) Note finally that these valence-bond expressions take exchange into account in a correct manner, because of the action of the projector Y.

B. Valence bond and the multipole expansion

Besides invoking the Rayleigh–Schrödinger second-order perturbation theory, London[16] and many workers after him also introduced an approximation to the interaction operator V^{AB} by expanding V^{AB} in a multipole series[49,50] and then truncating this series after the first few terms. Although we did not follow this procedure, but rather worked with the full interaction operator, we can still compare our results with those obtained in a multipole approximation by virtue of the high symmetry

(D_{2h}) of the ethylene monomer. In the multipole approximation one has to calculate expressions of the general form[50]

$$\frac{1}{R}\langle \psi_0^A \left| \frac{x_A^{n_1} y_A^{n_2} z_A^{n_3}}{R^{n_1+n_2+n_3}} \right| \psi_i^A \rangle \langle \psi_0^B \left| \frac{x_B^{m_1} y_B^{m_2} z_B^{m_3}}{R^{m_1+m_2+m_3}} \right| \psi_j^B \rangle \quad , \tag{10}$$

where R is the intermolecular distance, $x_A^{n_1} y_A^{n_2} z_A^{n_3}$ is a component of a 2^n-pole operator on A, $(n = n_1 + n_2 + n_3)$, and $x_B^{m_1} y_B^{m_2} z_B^{m_3}$ is a component of a 2^m-pole operator on B, $(m = m_1 + m_2 + m_3)$. Because the ground state of the ethylene monomer possesses A_{1g} symmetry, only locally excited states on A with the same symmetry as $x_A^{n_1} y_A^{n_2} z_A^{n_3}$, and locally excited states on B with the symmetry of $x_B^{m_1} y_B^{m_2} z_B^{m_3}$, will mix with this ground state. The symmetry of these operators is determined by their exponents n_1, n_2, n_3, and m_1, m_2, m_3 being even or odd. See Table I.

Conversely, using an untruncated operator V^{AB}, as we do in our calculations, and mixing only states of certain local symmetry with the ground state, Table I will tell us which terms of the multipole expansion are therewith implicitly taken into account. It is easy to see that the consecutive terms of V^{AB}, included by selecting the excited states of a certain symmetry, form a power series in $1/R^2$ in which the first term is the most important (around 10 bohr). Thus, by performing VB calculations on basis of states of well defined local symmetry, we have a means of comparing the exact results with those obtained from the multipole approximation. Furthermore, the analysis of the energy in terms of multipole interactions will guide us through the difficult process of selecting suitable states to be included in a multistructure VB calculation.

C. The nonorthogonality and spin problem

Although the valence-bond method has obvious conceptual advantages, its applications have not been very frequent. The main reason is the difficulty in the computation of the matrix elements

$$H_{ij} = \langle Y\psi_i^A \psi_j^B \left| H \right| Y\psi_i^A \psi_j^B \rangle \tag{11}$$

These matrix elements are hard to evaluate because of (1) the occurrence of the operator Y, (2) the nonvanishing overlap between the monomer states ψ_i^A and ψ_j^B. Much effort has been put into the elaboration of the first

TABLE I. Symmetry of multipole component under D_{2h}, n_1, n_2, n_3 = 0, 1, ···

$x^{2n_1} y^{2n_2} z^{2n_3}$	A_{1g}
$x^{2n_1+1} y^{2n_2+1} z^{2n_3}$	B_{1g}
$x^{2n_1+1} y^{2n_2} z^{2n_3+1}$	B_{2g}
$x^{2n_1} y^{2n_2+1} z^{2n_3+1}$	B_{3g}
$x^{2n_1+1} y^{2n_2+1} z^{2n_3+1}$	A_u
$x^{2n_1} y^{2n_2} z^{2n_3+1}$	B_{1u}
$x^{2n_1} y^{2n_2+1} z^{2n_3}$	B_{2u}
$x^{2n_1+1} y^{2n_2} z^{2n_3}$	B_{3u}

TABLE II. Effect of orthonormalization on second order energy. Results for geometry I (Fig. 1). Open shell orbitals π_A, π_B, π_A^*, π_B^*, σ_A^*, σ_B^*. Structures in both cases: $[\sigma\text{-core}]|\pi_A^2|[\pi_B^2]$, $[\sigma\text{-core}]|\pi_A\pi_B^*|[\pi_B\pi_B^*]$, $[\sigma\text{-core}]|\pi_A\sigma_A^*|[\pi_B\sigma_B^*]$. Here $[\sigma\text{-core}]$ stands for a 2s-electron structure with all 14 σ orbitals doubly occupied. The first order energy is not shown, since it is not affected by the orthonormalization.

Distance[a]	Nonorthogonal orbitals[b]	Orthogonal orbitals[b]
4 0	−337 10	−398 49
5 0	−166 73	−187 58
6 0	−71 57	−74 46
7 0	−28 17	−28 35
8 0	−11 62	−11 63
10 0	−2 64	−2 64
13 0	−0 49	−0 49
16 0	−0 13	−0 13

[a]Distance in bohr.[96] [b]Energy in 10^{-5} hartree.[96]

point;[51–57] but the second problem,[58–60] in some cases in connection with the first,[61,62] also received much attention. From the different ways of dealing with Y we have chosen Reeves' method based on the equivalent concept of "spin-bonded" functions.[63] We refer to the paper of Cooper and McWeeny[54] for the theoretical background of this method.

Reeves' algorithm assumes orthonormal orbitals, and so we are bound to orthonormalize. However, Slater[64] proved that this destroys chemical bond formation, and although it does not necessarily follow that orthonormalization is equally disastrous for van der Waals bonding, we still felt that we had to look into this problem.

In the first place, one notes that in the presence of closed-shell orbitals,[65] two orthonormalizations can be performed which do not alter the VB structures (except for a factor that disappears after normalization). The first orthonormalization regards the transformation of the closed-shell orbitals among themselves, and the second consists of the orthogonalization of the open-shell orbitals onto the closed shells. Further, we note that the orthonormalization of the open-shell orbitals among each other (by a matrix t) gives rise to a mixing of *all* structures with the same spin originating from the given open-shell orbital set. The mixing coefficients are elements of the tensor representation T(t).[66,67]

To study the effect of orthonormalization, we proceeded by constructing the matrix T(t) explicitly for the case of four open-shell electrons coupled to a singlet. The matrix t is obtained by a Gram–Schmidt orthonormalization[68] of the open-shell orbitals among themselves, all of which were first Gram–Schmidt orthonormalized onto the closed-shell orbitals which themselves were already Löwdin[68] orthonormalized among each other. Having obtained T(t), the secular problem is transformed to the original nonorthogonal orbital basis as described in Ref. 34. The results for a typical case with and without orthonormalization using the same VB structures are given in Table II.

As can be concluded from this table, the interaction

energy is slightly larger in the orthogonal case, which must probably be ascribed to the admixture of ionic VB structures by the orthogonalization. This contributes to the intramolecular correlation energy.[34] The effect of orthogonalization being very small indeed in the region of interest to us, we conclude that for the calculation of van der Waals forces, in contrast to chemical bonding, the use of orthogonalized monomer orbitals seems fully justified. In the actual calculation only the lowest two monomer MO's originating from the carbon 1s orbitals are kept doubly occupied. We have decided on the following orthonormalization procedure, which leaves the first order energy unaltered and the excited orbitals as close as possible to the original ones[69]: (1) Löwdin orthonormalize among each other all the orbitals that are occupied in the free monomers (seven σ-MO's and one π-MO); (2) Gram–Schmidt orthogonalize the excited orbitals onto all the thus obtained monomer orbitals; (3) Löwdin orthonormalize among each other the orthogonalized excited orbitals.

III. COMPUTATIONAL ASPECTS

A. Computer programs

The calculations were performed by the following chain of six programs:

(i) the integral program of IBMOL-5,[70] producing one and two electron integrals over contracted Cartesian Gaussian functions.[28] This program has been modified so that storage requirements are reduced by a factor of 2.6 and run times by a factor of 4 (for cases without symmetry). Still, this program is about 5 times slower than a later version of IBMOL[71] implemented after this work was finished;

(ii) the SCF program of IBMOL-5, which is essentially the same as that of ALCHEMY[72];

(iii) an orthonormalization program performing all the orthonormalizations described in this paper,

(iv) a transformation program[73] yielding integrals over MO's and based on an n^5 algorithm;[72,74]

(v) the first part of the VB program, which generates and sorts coefficients that do only depend on the spin, the number of orbitals, and the VB structures to be specified as input. Usually, these choices are kept unaltered in a series of related calculations (e.g., exponent or geometry variation), and in that case this step needs to be executed only once. The coefficients are generated by a FORTRAN translation of Reeves' algorithm,[63] published in ALGOL. Besides the coefficients occurring in the one electron part of the H matrix, the subroutine also calculates coefficients arising in the following expansion of the two electron part of the H-matrix elements:

$$\langle \psi_I \mid \sum_{\mu<\nu} \frac{1}{r_{\mu\nu}} \mid \psi_J \rangle = \sum_i \sum_j \sum_k \sum_l C(I,J;i,j,k,l)(ij \mid kl). \qquad (12)$$

This matrix element is based on VB structures constructed from orthonormal orbitals. It is one of the strong points of Reeves' algorithm that it does not put any constraints on the spin quantum number or on the number of singly occupied orbitals that can be handled. The summation in the right-hand side of Eq. (12) is restricted to a "canonical" order in $i\ j\ k\ l$.[75] The coefficients $C(I,J,i,j,k,l)$ permit the following factorization[76]:

$$C(I,J;i,j,k,l) = C_1(I,J)\,C_2(I,J;i,j,k,l) \qquad (13)$$

where $C_1(I,J)$ is independent of i,j,k,l and $C_2(I,J;i,j,k,l)$ can take on only 11 different values.[76] These properties of the C coefficients can be made useful to limit the input/output and main storage requirements of the VB program by the following procedure. Store $C_1(I,J)$ in canonical order of I and J discarding labels. Pack the value of $C_2(I,J;i,j,k,l)$ together with its six labels into one eight-byte word and perform all the ensuing processing on the $C_2(I,J;i,j,k,l)$ alone. After the construction of the H matrix is completed, the coefficients $C_1(I,J)$ are retrieved to multiply each H-matrix element.

The coefficients $C_2(I,J;i,j,k,l)$ are generated by Reeves' algorithm in a sequential order of I and J, with $I \geq J$, and in a rather arbitrary order of i,j,k,l. The four index transformation program, however, produces a canonically ordered list of integrals. In order to avoid enormous data transports at the execution of Eq. (12), the generation of the $C_2(I,J;i,j,k,l)$ is followed by a sorting procedure which brings the coefficients in canonical order of i,j,k,l. Because we followed closely an idea of Yoshimine[77] to minimize the number of I/O operations, here as well as in the construction of the H matrix, we refer to that author for more details.[78]

(vi) The second part of the VB program constructs the S and H matrix [Eq. (11)] and transforms these matrices optionally to nonorthogonal orbitals by means of a matrix T(t). We intend to give more details about the actual construction of T(t) in a later paper. As the last step the secular problem is solved, employing fully the blocked structure of S (only structures with the same orbital occupancy give an overlap). The present version of the program contains a diagonalization subroutine based on the Givens–Householder–QR scheme,[79,80] demanding core space for one full matrix and a few columns. At the moment this is the only, but easily removable, bound to the maximum number of structures.

We conclude by quoting a few representative figures, all regarding double precision arithmetic on an IBM S370/158 computer: The computation of the integrals for the ethylene dimer in the 84-dimensional primitive basis specified below took on the average 54.4 min. The transformation of a list of integrals derived from 52 AO's to one on basis of 28 MO's needed 32 min. The generation of a list of 70 606 coefficients deriving from 183 singlet structures, 28 orbitals, and 32 electrons takes 79 sec. The reordering of this list takes 19 sec. The construction of the corresponding 183-dimensional H matrix takes 24 sec and the solution of the secular problem 263 sec.

TABLE III AO basis for ethylene

		Coefficients[a]	Exponents[b]
C(6,3/3,2)	s	0 026916	1087 10
	s	0 197505	163 867
	s	0 845777	37 4090
	s	0 574338	10 5181
	s	0 482412	3 32078
	s	1 0	0 2934J6
	p	0 212696	4 20169
	p	0 878780	0 858491
	p	1 0	0 202063
H(3/2)	s	0 143689	6 4805
	s	0 913002	0 98104
	s	1 0	0 21798

[a]This work [b]References 81 and 82

B. Atomic and molecular orbitals

For reasons of economy, a rather small basis of contracted GTO's was employed a C(6, 3/3, 2) set on the carbons and a H(3/2) set on the hydrogens. See Table III The hydrogen exponents taken from Ref. 81 were scaled by a factor 1.2 The contraction coefficients of all AO's were obtained from a calculation on the free ethylene in an uncontracted basis, applying the rules of Dunning.[83] In accordance with the findings of Moskowitz and co-workers,[66] an isotropic basis of p orbitals was observed to give good results, so we used the same exponents and contractions for the three p orbitals. Several degrees of contraction have been tried, with the rather loose scheme which was finally decided on (Table III) yielding the most acceptable deviation from the uncontracted computation. For comparison's sake several SCF results of ethylene are listed in Table IV Quadrupole moments are given in Table V.

No atomic polarization functions (carbon $3d$, hydrogen $2p$) were included because this would have lead to prohibitive calculation times. Using the faster integral program available now, some tests are underway to study the effect of such functions. The AO basis used gives rise to 26 molecular orbitals on each monomer, of which only the lowest 8 are doubly occupied in the ground state, the virtual ones playing the role of molecular polarization functions. To keep the VB calculations tractable, we were forced to make a selection of the virtual orbitals to participate in the dimer VB calculations. From earlier work,[34] it could be inferred that an optimization of the virtual orbitals is important for the correct estimate of induction and dispersion energy. Therefore, it was decided not to proceed simply with the virtual orbitals originating from a ground state Hartree–Fock calculation, but to determine individually each MO to be included in the pool of excited MO's. To this end a π-electron was consecutively promoted to the lowest orbital of each occurring symmetry: b_{3g}, b_{3u}, a_{1g}, b_{2u}, b_{1g}, and b_{1u}, whereupon these six orbitals were one by one optimized through open-shell SCF procedures on the respective excited states. The orbitals obtained this way are orthogonal neither on each other nor on the ground state orbitals (except perhaps for symmetry reasons), which is why we performed a preliminary Gram–Schmidt orthonormalization on the separate set of monomer orbitals, preserving the energy order of the orbitals. The orbital energies and symmetries are shown in Table VI.

TABLE IV Ethylene calculations in different basis sets

Basis	Energy (hartrees)
STO-4G[a]	−77 85810
C(6,3/3,2), H(3/2)[b]	−77 900063
C(6,3), H(3)[b]	−77 901246
$s+p$ limit[c]	−78 0062
Hartree–Fock limit[d]	−78 0623

[a]Reference 84
[b]This work, atomic coordinates from Ref 85
[c]Reference 86
[d]Reference 87

TABLE V Quadrupole moments,[a] ethylene (a u)

	Q_{xx}	Q_{yy}	Q_{zz}
C(6,3/3,2), H(1/2)[b]	1 1533	1 5875	−2 7408
Double zeta[c]	1 4642	1 4982	−2 9624
Experimental[d,e]	...	..	−2 98[d]
			−2 75[e]

[a]$Q_{\alpha\alpha} = (3\langle r_\alpha^2\rangle - \langle r^2\rangle)/2$
[b]This work, coordinates of monomer A, Fig 1
[c]Reference 88
[d]Reference 89
[e]Reference 90

C VB structures

As was pointed out in the foregoing section, only those VB structures are to be taken into account that mix under the first few multipole operators resulting from the expansion of the interaction operator. Only single excitations on each of the monomers have to be included to calculate a substantial part of the London–van der Waals

TABLE VI Ethylene orbitals included in VB calculations, 1–8 from closed-shell SCF on ground state, 9–14 from open-shell SCF on singly excited states

nr	Symmetries[a]	Energy (hartrees)
1	a_{1g}	−11 287616
2	b_{2u}	−11 266116
3	a_{1g}	−1 025231
4	b_{2u}	−0 783828
5	b_{3u}	−0 641625
6	a_{1g}	−0 582004
7	b_{1g}	−0 499021
8	b_{1u}	−0 375118
9	b_{3g}	0 007255
10	b_{3u}	0 060831
11	a_{1g}	0 082727
12	b_{2u}	0 132136
13	b_{1g}	0 225237
14	b_{1u}	1 062311

[a]Point group D_{2h}

TABLE VII Contributions[a] of local excitations to the lowest multipole operators of certain symmetry

Dipole moment		Quadrupole moment	
	5 → 14[b]		5 → 11
B_{1u}	6 → 14		6 → 11
(z)	8 → 11	A_{1g}	7 → 13
	4 → 9	(x^2, y^2, z^2)	8 → 14
	3 → 12		4 → 12
	6 → 12		5 → 10
B_{2u}	4 → 11		3 → 13
(y)	8 → 9		6 → 13
	7 → 10	B_{1g}	7 → 11
	5 → 13	(xy)	4 → 10
	3 → 10		5 → 12
	6 → 10		7 → 9
B_{3u}	5 → 11	B_{2g}	5 → 14
(x)	7 → 12	(xz)	8 → 10
	4 → 13		3 → 9
			6 → 9
		B_{3g}	4 → 14
		(yz)	8 → 12

[a]Table relates to monomer A in Fig. 1. For monomer B in geometry II, the y and z coordinates must be interchanged.
[b]Orbital numbers from Table VI

energy. Induction energy is obtained by exciting one molecule only, whereas the simultaneous excitations of both monomers yield dispersion forces. The subsystems can be excited to a singlet or a triplet state, and since two singlets as well as two triplets can couple to a singlet dimer state, each pair of local excitations will yield two linearly independent structures. In Table VII the excitations are given which admix to the ground state under the various components of the dipole and the quadrupole operator. In this table, as in the calculations, the assumption is made that the carbon 1s electrons do not contribute to the induction and dispersion energy, and consequently the lowest two monomer MO's are kept doubly occupied. From Table VII one derives Table VIII, where the numbers of VB structures are exhibited which must be included to get a full description of the following second order interactions: induced dipole/induced dipole, induced dipole/induced quadrupole, induced quadrupole/induced quadrupole, and permanent quadrupole/induced dipole. At this point it must be stressed again that by using a nontruncated interaction operator, symmetry arguments can not completely separate the higher multipole interactions from the lower ones. For instance, the result to be presented as "z-component part of dipole/dipole dispersion energy" includes in fact also the z-component part of dipole/octupole, octupole/octupole, etc. dispersion energy. From Tables I and VII it can immediately be deduced which of the higher multipole interactions are also included under the different choices of local symmetry. From these tables it can also be concluded, even though we did not include atomic polarization functions, that all symmetries are represented except one: A_u. The lowest order operator with A_u symmetry is the xyz component of the octupole operator which couples with the xy component of the quadrupole operator on the other center, giving rise to a R^{-12}-dependent contribution to the dispersion energy.

No ionic structures were added because the included covalent VB structures should account for the electrostatic, induction, and dispersion forces (Sec. II. A). The covalent structures are also expected to describe exchange forces rather well although we use orthogonalized orbitals, because the first order interaction between closed-shell monomers is invariant under orthogonalization (Sec. II. C). So, the first order exchange forces are accounted for exactly: only for short distances are ionic structures expected to improve the exchange energy (in second order). Moreover, the inclusion of ionic structures would greatly complicate the formalism, since they are known to yield a distance dependent contribution to the intramolecular correlation energy [34] implying that we would have to take the fully correlated monomers as the reference system.

IV. RESULTS

To date, almost all semiempirical calculations on the conformation of biopolymers and molecular crystals apply the pair approximation neglecting three and more body interactions.[91] A recent study[41] looked into the pair approximation for the case of the ethylene crystal and concluded that three body forces are indeed negligible. So, in accordance with the purpose of this work, we restrict the attention to the dimer.

The interaction energy as a function of distance has

TABLE VIII Number of VB structures contributing to the components of second order multipole interactions

Multipole interaction	Geometry I	Geometry II
$\mu_z^A-\mu_z^B$	32	48
$\mu_y^A-\mu_y^B$	72	48
$\mu_x^A-\mu_x^B$	50	50
Total dipole–dipole (dispersion)	154	146
$\mu_z^A-Q_{yz}^B+Q_{yz}^A-\mu_z^B$	96	120
$\mu_y^A-Q_{yz}^B+Q_{yz}^A-\mu_y^B$	96	80
$\mu_x^A-Q_{xz}^B+Q_{xz}^A-\mu_x^B$	60	80
Total dipole–quadrupole (dispersion)	252	280
$Q_{yz}^A-Q_{yz}^B$	72	72
$Q_{xy}^A-Q_{xy}^B$	50	30
$Q_{xz}^A-Q_{xz}^B$	18	30
$Q_{yz}^A-Q_{yz}^B$	32	32
Total quadrupole–quadrupole (dispersion)	172	164
Permanent quadrupole–induced dipole (induction)	8	10
Total multipole interaction (up to quadrupole–quadrupole)	586	600

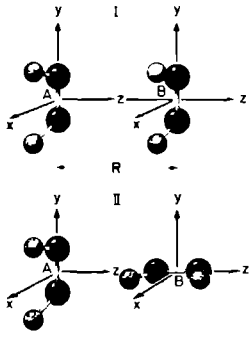


FIG. 1. Geometries of the ethylene dimer considered in this work.

been computed for two relative orientations of the ethylene molecules (Fig. 1).

(i) Geometry I is the structure occurring in the stacks of π-electron systems mentioned in the introduction. This structure also constitutes the simplest possible model exhibiting all the basic features of transannular interaction, such as it appears for instance in the (m,n)-paracyclophanes or bicyclooctatriene.[92]

(ii) Geometry II is the structure yielding the maximum electrostatic quadrupole–quadrupole attraction.[50] [Note that the quadrupole moment of ethylene is close to that of a linear molecule (Table V), with the axis perpendicular to the molecular plane.]

The importance of quadrupole–quadrupole interaction for determining the spatial arrangement of molecular crystals has often been emphasized,[93–95] whereas in other work[12,13] the significance of this interaction in solids is cast into doubt. So it is a point of interest to compute the size of this term.

In Table IX, the total first- and second order interaction energies defined in Sec. II are given; the same results are graphically presented in Figs. 2 and 3. The

TABLE IX. First order, second order, and total interaction energies. Energy units 10^{-5} hartree.[96] Distance R between the centers of mass in bohr.[96] Geometries given in Fig. 1. Definitions of interaction energies given by Eqs. (8) and (9). Zero point of energy: -155.8001251 hartree.

	Geometry I			Geometry II		
R	$\Delta E_{VB}^{(1)}$	$\Delta E_{VB}^{(2)}$	ΔE_{VB}	$\Delta E_{VB}^{(1)}$	$\Delta E_{VB}^{(2)}$	ΔE_{VB}
8.0	8104.39	−3532.38	4572.01	*	*	*
8.5	4234.*	−1656.**	2577.**	**	*	*
9.0	2252.37	−961.26	1291.11	4807.46	−236.45	4571.01
9.5	1253.55	−617.96	635.61	864.22	−102.36	761.86
10.0	702.64	−373.88	328.76	104.71	−46.01	58.70
9.0				−8.05	−21.75	−29.80
10.0	5.47	−8.29	−2.82	−13.44	−11.07	−26.51
11.0				−10.12	−5.97	−16.09
11.0	9.96	1.76	5.16	−5.53	−2.02	−7.55
12.0	1.75	−0.50	1.25	−2.14	−0.57	−2.71

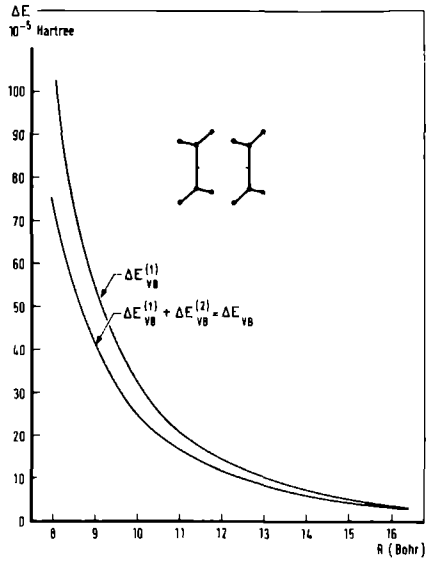


FIG. 2. First order and total valence-bond interaction energies for Geometry I (Table IX).

first order energy consists in general of short range exchange and penetration effects on the one hand, and long range electrostatic interactions on the other. In Table

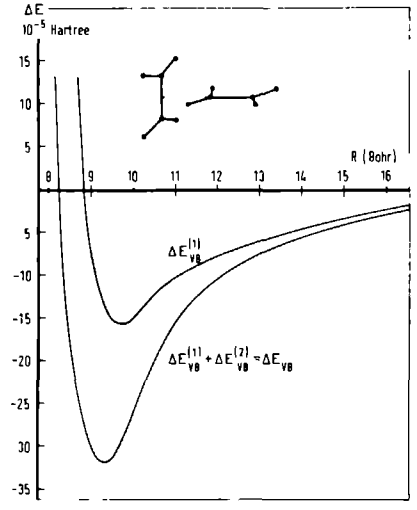


FIG. 3. First order and total valence-bond interaction energies for Geometry II (Table IX).

TABLE X. Decomposition first order energy into electrostatic quadrupole–quadrupole $\Delta E^{(1)}_{Q-Q}$, quadrupole–hexadecupole $\Delta F^{(1)}_{Q-H}$, hexadecupole–hexadecupole $\Delta E^{(1)}_{H-H}$ interaction energy and short range exchange and penetration effects. Units 10^{-5} hartree.

	Geometry I				Geometry II			
R (bohr)	$\Delta E^{(1)}_{Q-Q}$	$\Delta E^{(1)}_{Q-H}$	$\Delta F^{(1)}_{H-H}$	Exchange + penetration	$\Delta F^{(1)}_{Q-Q}$	$\Delta E^{(1)}_{Q-H}$	$\Delta E^{(1)}_{H-H}$	Exchange + penetration
4 0	4407 69	−10312 75	11843 05	12166 60	−2604 47	6198 12	−5030 01	...
5 0	1444 31	−2162 74	1589 55	3374 69	−853 43	1299 84	−675 12	...
6 0	580 44	−603 57	308 06	767 64	−342 97	362 76	−130 84	4918 51
7 0	268 55	−205 17	76 93	145 24	−158 68	123 31	−32 68	932 27
8 0	137 74	−80 57	23 12	22 34	−81 39	48 42	−9 82	147 50
9 0	76 44	−35 33	8 01	...	−45 16	21 23	−3 40	19 28
10 0	45 13	−16 89	3 10	1 33	−26 67	10 15	−1 32	2 40
11 0	28 03	−8 67	1 32	...	−16 56	5 21	−0 56	1 79
13 0	12 16	−2 69	0 29	0 16	−7 18	1 62	−0 12	0 15
16 0	4 30	−0 63	0 05	0 02	−2 54	0 38	−0 02	0 04

X and Figs. 4 and 5, these long and short range contributions to the first order energy are separated. The electrostatic force has been split as follows: the classical quadrupole-quadrupole interaction $\Delta E^{(1)}_{Q-Q}$ has been computed using our own quadrupole moments shown in Table V. Lacking a computer program generating fourth moments, we took hexadecupole values from Ref. 88, which are of double-ζ quality. We employed these for the hexadecupole–quadrupole interaction $\Delta E^{(1)}_{Q-H}$ and hexadecupole–hexadecupole interaction $\Delta E^{(1)}_{H-H}$. Invoking Rose's formula,[49] one easily derives the values given in Table X. The difference between the quantum mechanical first order energy and the classical electrostatic interaction (up to hexadecupole–hexadecupole) is also given in Table X under the heading "exchange and penetration energy." Besides the real exchange and penetration effects, these values also contain the deviations caused by the use of hexadecupole moments that are not derived from our monomer wavefunction and the contribution of higher multipole interactions. Work is in progress to refine the analysis on these points.

In Table XI, the second order energy is decomposed into the various components of the multipole moments, a decomposition that is not obtained by expanding the interaction operator but is induced by the use of local symmetry, as described in Sec. II. Making this decomposition, one implicitly assumes that the contributions to the second order in our definition (7) are

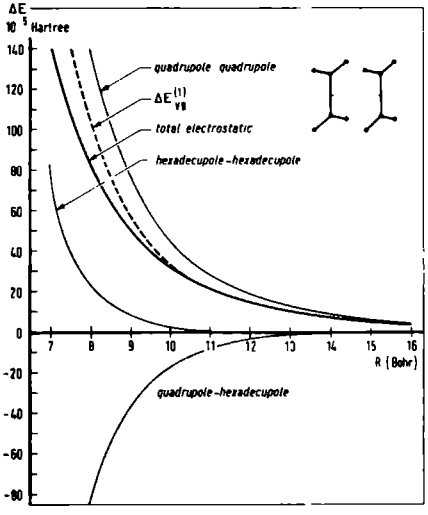


FIG. 4. Decomposition of first order energy for Geometry I (Table X).

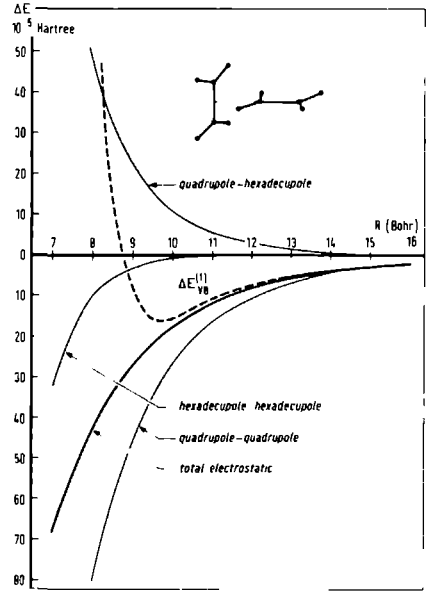


FIG. 5. Decomposition of first order energy for Geometry II (Table X).

TABLE XI. Decomposition of second order energy into components of dipole–dipole, dipole–quadrupole, and quadrupole–quadrupole dispersion and permanent quadrupole-induced dipole induction. Energy in 10^{-5} hartree,[96] distance in bohr.[96]

R	$z-z$	$y-y$	$x-x$	$z-y^2$ $+y^2-z$	$y-xz$ $+xz-y$	$y-xy$ $+xz-x$	y^2-x^2	$xy-xy$	$xz-xz$	$yz-yz$	Induction
					Geometry I						
4.0	−20.31	−469.59	−95.64	−47 05	−125 09	−14.91	−52 78	−16 76	−2 32	−14 21	−168.92
5.0	−10 33	−225 82	−44.43	−16 41	−51.20	−5 83	−16 12	−5 29	−0 63	−4.72	−9 80
6.0	−4.84	−95.99	−21.10	−6 81	−18 15	−2 23	−5 43	−1 68	−0.20	−1.32	−3.62
7.0	−2 34	−40.31	−10 64	−2 96	−6 27	−0.89	−1 92	−0 58	−0.06	−0.38	−1.63
8.0	−1 20	−18 29	−5.68	−1 30	−2 33	−0.38	−0 69	−0.21	−0 02	−0.12	−0.69
10.0	−0 37	−4.89	−1 67	−0 29	−0 43	−0 08	−0.10	−0.03	−0.00	−0.01	−0 15
13.0	−0 09	−1 05	−0 47	−0 04	−0 06	−0.01	−0.01	−0 00	−0.00	−0.00	−0.02
16.0	−0.02	−0.30	−0.15	−0.00	−0 01	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00
					Geometry II						
6.0	−28 15	−32 62	−56.00	−30 16	−21 91	−32 45	−5 09	−0 44	−3.81	−2.73	−24.36
7.0	−15.15	−13.03	−27.18	−12 53	−7 70	−14.57	−1.29	−0 14	−1 60	−0.84	−8.45
8.0	−8.16	−5.26	−13.15	−5.50	−2 60	−6 19	−0.38	−0 04	−0 56	−0.25	−3.93
9.0	−4.49	−2 33	−6.64	−2 48	−0 94	−2 70	−0 14	−0.01	−0.18	−0.07	−1.74
10.0	−2.57	−1 15	−3 55	−1.18	−0 39	−1 26	−0 06	−0.01	−0.07	−0.03	−0.82
11.0	−1.50	−0 36	−2 28	−0 58	−0 18	−0 62	−0.02	−0.00	−0.03	−0.01	−0.40
13.0	−0.59	−0.20	−0.72	−0.17	−0 04	−0 17	−0.00	−0.00	−0.00	−0.00	−0.11
16 0	−0 19	−0.06	−0 21	−0 04	−0 01	−0.04	−0 00	−0.00	−0.00	−0.00	−0.03

additive,' just as they are in the corresponding second order perturbation energy. To verify this assumption we performed four different sets of calculations: one simultaneously including all structures that contribute to the $\vec{\mu}-\vec{\mu}$ (dipole–dipole) part of the dispersion; one calculation giving the full $\vec{Q}-\vec{\mu}$ (quadrupole–dipole) dispersion; one yielding the full $\vec{Q}-\vec{Q}$ (quadrupole–quadrupole) dispersion, and a computation on the $\vec{\mu}-\vec{\mu}$ plus $\vec{Q}-\vec{\mu}$ second order energy. A simplification could be made since it appeared from the calculations on the individual components that almost all VB structures representing the coupling of triplet excited monomers give rise to negligible contributions (in general less than 10^{-6} hartree; exceptions are discussed in the next section). The computations were accordingly done with omission of all noncontributing triplet–triplet structures. The differences between the energies of the four overall calculations and the sum of the corresponding component energies are very small, less than 0.6×10^{-6} hartree for $R=6$ bohr, less than 10^{-7} hartree for $R=8$ bohr.

The ratios of the different terms in the multipole expansion to the second order energy are plotted in Figs. 6 and 7, clearly exhibiting the convergence of the multipole series. In Fig. 6, the corresponding ratio of the contribution of the $\pi-\pi^*$ excitation has also been drawn. Note that this part of the dipole–dipole dispersion is obtained from a three-structure VB calculation on the basis of only the ground state and the two states representing the intermolecular coupling of the $T(\pi\pi^*)$ and $V(\pi\pi^*)$ states, respectively, whereas Table VIII tells us that the total second order energy is obtained from as many as 600 VB structures. The effect of the $\pi-\pi^*$ transition is not shown in Fig. 7, because no special contribution is noticeable in the case of Geometry II. In neither of the two geometries does any other structure yield a dominant contribution.

Wondering if possibly the first order and induction energy could be accounted for by the SCF formalism, we performed two sets of LCAO-MO-SCF calculations on the dimer, treating it as if it were one "supermolecule." One set was done with the CNDO/2 program,[97] using the original parameter setting,[96] and the other calculations were of the *ab initio* type, employing the GTO basis given in Table III. The results are given in Figs. 8 and 9.

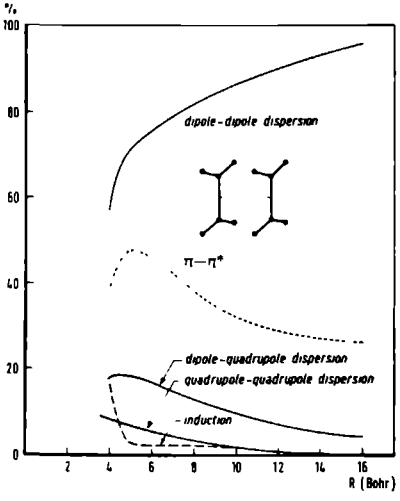


FIG. 6. Relative contributions to second order energy for Geometry I. Absolute values given in Table XI.

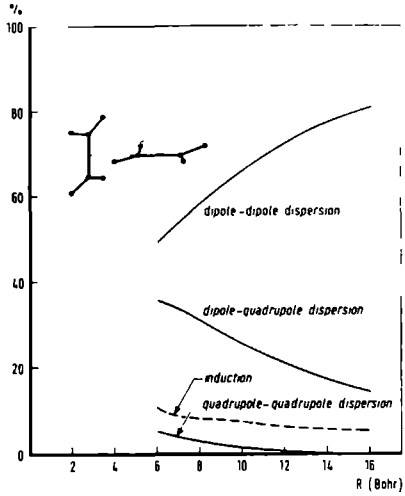


FIG 7 Relative contributions to second order energy for Geometry II Absolute values given in Table XI

V. DISCUSSION

A. Long range interactions and their multipole expansions

The difference between the interaction of two parallel (Fig. 2) and two perpendicular (Fig. 3) ethylene molecules, one being repulsive and the other attractive, is seen from Tables IX and X to be mainly due to the first order electrostatic interaction. At large and intermediate distances, the $\overline{Q}$–$\overline{Q}$ (quadrupole–quadrupole) interaction is the dominant term in this first order energy, but at smaller distances the higher multipoles become important as well. Note that the multipole expansion in first order diverges completely at $R \approx 5$ bohr, with even $\Delta E^{(1)}_{Q-Q} < \Delta E^{(1)}_{M-M}$ for Geometry I. Clearly, penetration effects have rendered the expansion invalid in this region. Note also that the three term expansion ceases to be an adequate representation of the intermolecular potential much earlier for distances smaller than 10 bohr, the interaction between higher moments than hexadecupoles must be included. The decisive role of the $\overline{Q}$–$\overline{Q}$ interaction in determining the energy difference between the two geometries considered suggests that these interactions are also important in determining the structure of molecular crystals.[93–95] In a crystallographic environment, the effect of the surrounding quadrupoles on a molecule may average out though.[12,13] Still, our calculations show that first order electrostatic forces cannot be neglected *a priori*. If the $\overline{Q}$–$\overline{Q}$ interaction is calculated, higher multipoles cannot consistently be omitted.

As far as the second order forces are concerned they are always attractive and, therefore, they will certainly contribute to the cohesion energy of the crystal. The attraction is slightly larger for the perpendicular geometry than for the parallel one Comparing Figs 6 and 7, we note that the dipole–dipole dispersion has relatively more weight in the parallel geometry. In Table XI, the source of this phenomenon can readily be located it is the ν–ν component of the dispersion, which in Geometry I receives a considerable contribution from the π–π^* transition on both monomers In Geometry II, these π–π^* transitions contribute to different dipole–dipole components, and it appears that they play no particular role in this geometry. In the second order energy, as in the first order, we find that the leading term in the multipole expansion does not suffice for a good description of the interaction energy, so that we should also take higher multipoles into account. The induction energy is not very large, which is not a widely accepted fact. A point of criticism could here be that the induction of a quadrupole by the permanent quadrupole on the other monomer has not been calculated. But since the corresponding energy has a R^{-10} dependence and since the quadrupole-induced dipole (R^{-8}) term is already quite small, we felt safe in neglecting it

B Penetration and exchange effects

Although we have not calculated these effects directly, their occurrence can easily be recognized from our results. Figure 6 shows a sudden drop in the relative dipole–dipole contribution originating from a steep rise in the induction energy, which we ascribe to the same penetration of the charge distribution that also causes the divergence of the expanded first order energy. If this assignment is correct, the exponential increase of the first order repulsion at about 7 and 8 bohr for Geometry I and II, respectively, must be caused by exchange

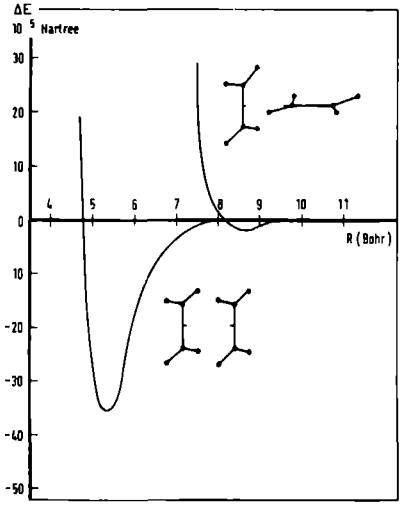


FIG. 8 Interaction energies from CNDO calculations on the supermolecule.

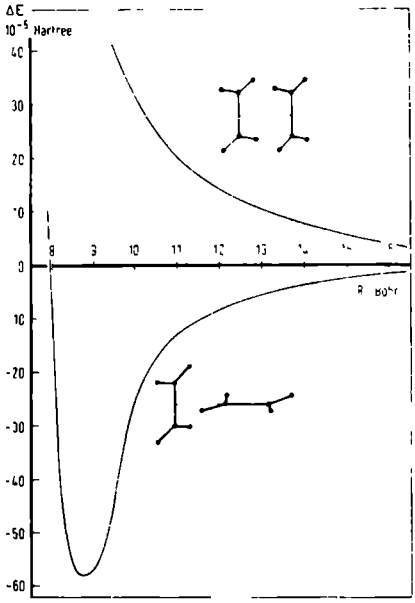


FIG. 9. Interaction energies from *ab initio* LCAO-SCF calculations on the supermolecule

effects. Actually, the difference of about 1 bohr in the "hard core" diameters for Geometry I and II, together with a scattering diameter σ of 8.3 bohr for Geometry II (Fig. 3) matches a geometrical model that attributes a van der Waals radius of 3.7 bohr to a spherical methylene group. Pauling[99] suggests 3.8 bohr for this radius a number which is also in accordance with the C ··· C and C ··· H contact distances proposed by Kitaigorodski[13]

The over-all repulsive interaction between two parallel ethylene molecules is also in agreement with the well-known fact that a strong steric hindrance exists between two transannular ethylenic moieties. This has been discussed earlier in connection with the influence on the $\pi-\pi^*$ transitions localized in the double bonds[100] and it has been investigated by means of extended Hückel calculations.[92]

C. Comparison with experiment

Figure 3, relating to the perpendicular geometry, shows a van der Waals well depth ϵ of 33.5×10^{-5} hartree at $R_0 = 9.4$ bohr and a scattering diameter σ of about 8.3 bohr. The first order contribution to ϵ is 12.5×10^{-5} hartree.

Experimental values for the van der Waals well of the ethylene dimer are as follows: $\epsilon = 72.9\times10^{-5}$ hartree, $\sigma = 7.7$ bohr[101]; or $\epsilon = 65.0\times10^{-5}$ hartree, $\sigma = 8.0$ bohr.[102] While our scattering diameter agrees reasonably with both values, the discrepancy between the measured and computed values of ϵ is considerable. Since the experimental values have been obtained from viscosity data in which a certain rotational averaging is included, the difference may be even larger. Two points are relevant in this respect. Earlier calculations[34] on He_2 pointed to a need for the optimization of the excited monomer orbitals and also showed that the excited state SCF method is not an adequate tool to this end. (The behavior of virtual SCF orbitals is worse.) So, not having employed excited orbitals which are fully optimized for a description of the long range interaction, and not having included atomic polarization functions, we feel that our results may underestimate the dispersion energy. Also the question whether Rydberg orbitals should have been included seems appropriate, since the ethylene spectrum shows several Rydberg series and there is even some doubt if the π^* orbital may not be Rydberg-like[103] One must realize, however that the *ab initio* calculation of van der Waals forces between molecules of this size is not an easy matter and requires very time-consuming computations. In view of this, we feel that our results are promising. Further work on the optimization of excited orbitals, including atomic polarization functions, is in progress

Another reason for the disagreement with the experimental results may lie on the experimental side. The viscosity data were interpreted in an isotropic 6-12 potential, which is not very realistic for this case. To give an impression how sensitively the ϵ *values depend* on the measured viscosities, we quote Flynn and Thodos,[101] who call the difference "plausible" between viscosity ϵ/k values of 410 °K and 208 °K found for *n*-butane by themselves and Hirschfelder *et al.*[102] respectively. This underlines the unreliability of the few experimental data available to gauge the parametrization of semiempirical calculations

Even though the second order energy may be underestimated we still feel that the ratios of the different contributions (Figs. 6 and 7) are of correct magnitude because they are computed in a single consistent manner without introduction of any *a priori* prejudices about their importance.

D. Interactions between stacked π systems

The repulsive interaction calculated between two parallel ethylene molecules may seem in contradiction to the opinion[1-4] that the stabilization of the helical conformation of DNA is mainly caused by attractive vertical interactions between the bases and, equally to the idea that $\pi-\pi$ charge transfer complexes are stabilized by van der Waals interactions. Although the ethylene dimer evidently falls short in exhibiting all the properties of interactions between DNA bases or large charge transfer complexes we can still offer a possible explanation on the basis of Fig. 6 Here a very pronounced contribution of the π electrons to the dispersion energy is observed, and we may therefore, not without justification, speculate that in the case of large, very polarizable π systems, the second order energy is able to surmount the first order repulsion, that is, of course, in the region where the Pauli repulsion is still negligible. This

outstanding contribution of the $\pi-\pi^*$ excitations is only found for parallel π systems.

E. Local symmetry, additivity, and perturbation theory

Two important conclusions concerning the method of calculation of the second order energy can be drawn. The triplet–triplet VB structures hardly mixing with the ground state, it appears that local spin selection rules, forbidding the mixing of triplet–triplet and singlet–singlet structures, are very well preserved upon formation of the dimer. So, VB calculations on the weak interactions considered in this work can be drastically simplified by omitting all the structures representing the triplet–triplet coupling. There is a small exception though: for the short distances of the parallel geometry, 11 triplet–triplet structures, all belonging to the $\nu-\nu$ dispersion component, contribute slightly to second order energy (3.09×10^{-5} and 1.66×10^{-5} hartree for R equal to 4.0 and 5.0 bohr, respectively). The most noticeable among these 11 functions is the one representing the coupling of the $T(\pi\pi^*)$ states. Subsequent inclusion of all triplet–triplet structures gave no further improvement, and it can thus be concluded that down to 4.0 bohr the triplet–triplet couplings give rise to negligible contributions to the ground state.

Another important conclusion to be drawn regards the possibility of approximating the lowest eigenvalue of the secular problem over VB structures by a perturbationlike formula. Recall that each component of the second order energy has been computed on a basis of VB structures which are adapted to the local symmetry, in this case characterized by $D_{2h}\otimes D_{2h}$. The off-diagonal H-matrix elements connecting blocks of different local symmetry contain only terms arising from the interaction operator V^{AB}. The high degree of additivity in the multipole components of the second order energy shows these elements to be so small that the higher order terms in formula (5) can be neglected between structures of different local symmetry, thus enabling a componentwise construction and diagonalization of H. This in itself is already a great help in keeping the method tractable for large complexes, but it also points to a further potential simplification. Although from our present calculations it cannot be inferred with absolute certainty that the V^{AB} terms within the symmetry blocks are equally small, there is no reason why they should not be. This additivity even holds in the region where the Pauli repulsion has become large, and we may therefore tentatively conclude that a perturbationlike second order formula may be applicable to the lowest eigenvalue of the H matrix including exchange for the whole range of the potential curve. It should be said that the H matrix in this conclusion is assumed to be over a basis of the eigenvectors of H^A and H^B, as described in Sec. II of this paper, because otherwise H contains nondiagonal terms originating from H^A or H^B.

F. SCF Results

Let us finish this section by making a few comments on the SCF results presented in Figs. 8 and 9. It is noteworthy that the CNDO calculations predict the parallel geometry to be the more stable one, whereas the *ab initio* results fall in line with the VB predictions in this respect. A similar disagreement of CNDO with *ab initio* SCF has been noted before for $(HCN)_2$,[104] and has there been ascribed to the neglect of three and four center repulsions causing CNDO to favor a cyclic structure. This explanation being very plausible, it makes the CNDO results for this complex meaningless.

The *ab initio* results on the other hand, seem surprisingly good, with a σ value of 8.0 bohr and $\epsilon = 57.6\times10^{-5}$ hartree, which may be compared with the experimental values[102] σ – 8.0 bohr, $\epsilon = 65.0\times10^{-5}$ hartree. However, as a check on the usefulness of these results, we performed an SCF computation on the free monomer A in its own AO basis augmented by the vacant AO basis of monomer B placed at a distance 9.0 bohr, assuming Geometry II. This basis set enlargement gave an energy improvement of 56.09×10^{-5} hartree, which might become somewhat smaller by accounting for the filling of the orbitals on monomer B. Still, this proves that the splendid SCF curve is partly due to the mathematical artifact of distance dependent basis set enlargement, and has little physical significance. It is difficult to separate the physical interaction energy from these SCF results, since the energy lowering by the basis set enlargement is a nonadditive effect. This pitfall, threatening those who apply small basis SCF to the computation of intermolecular forces was first noted by Kestner[30] in a discussion of early *ab initio* SCF calculations on He_2.[29] Our results bear witness again to the fact that calculations on van der Waals interactions employing the SCF "super-molecule" approach must necessarily be looked upon with mistrust as long as the monomer bases do not approach the Hartree–Fock limit.

ACKNOWLEDGMENT

We would like to thank Dr. C. Moser for his hospitality at CECAM, where part of the programs have been developed, A. van Berkel for his help in writing the program for nonorthogonal orbitals, and Dr. F. B. van Duijneveldt and Dr. M. van Hemert for valuable discussions.

*Supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

[1] *Molecular Associations in Biology*, edited by B. Pullman (Academic, New York, 1968), especially Pullman's introduction.
[2] H. de Voe and I. Tinoco, Jr., J. Mol. Biol. **4**, 500 (1962).
[3] P. Claverie, B. Pullman, and J. Caillet, J. Theor. Biol. **12**, 419 (1966).
[4] P. Rein and M. Pollak, J. Chem. Phys. **47**, 2039 (1967).
[5] A. Saran, D. Perahia, and B. Pullman, Theor. Chim. Acta **30**, 31 (1973), and references therein.
[6] C. K. Prout and J. D. Wright, Angew. Chem. Int. Ed. Engl. **7**, 659 (1968).
[7] M. J. S. Dewar and C. C. Thompson Jr., Tetrahedron Suppl. **7**, 97 (1966).
[8] R. J. W. Le Fèvre, D. V. Radford, and P. J. Stiles, J. Chem. Soc. B **1968**, 1297.
[9] M. J. Mantione, Theor. Chim. Acta **15**, 141 (1969).
[10] M. W. Hanna and J. L. Lippert, in *Molecular Complexes*, edited by R. Foster (Paul Elek Scientific Books, London, 1973).
[11] R. E. Hancock and J. N. Murrell, Chem. Soc. Faraday

Trans II 1, 115 (1973)
[12]D P Craig, R Mason P Pauling and D P Santry Proc R Soc Lond A 286, 98 (1965)
[13]A I Kitaigorodsky, *Molecular Crystals and Molecules* (Academic, New York, 1973)
[14]D E Williams J Chem Phys 45, 3770 (1966)
[15]A I M Rae, Mol Phys 16, 257 (1969)
[16]F London, Z Phys 63, 245 (1930)
[17]F London J Phys Chem 46, 305 (1942)
[18]A van der Avoird J Chem Phys 47, 3649 (1967)
[19]D M Chipman, J D Bowman, and J O Hirschfelder J Chem Phys 59, 2830 (1973), and references therein
[20]A van der Avoird Chem Phys Lett 1, 429 (1967)
[21]J N Murrell, M Randić, and D R Williams Proc R Soc Lond A 284, 566 (1965)
[22]J L Lippert, M W Hanna, and P J Trotter, J Am Chem Soc 91, 4035 (1969)
[23]P Claverie, in Ref 1
[24]G R Pack, H Wang, and R Rein, Chem Phys Lett 17, 381 (1972)
[25]R Bonaccorsi, R Cimiraglia E Scrocco, and J Tomasi, Theor Chim Acta 33, 97 (1974)
[26]A Riera and W J Meath, Int. J Quantum Chem 7, 959 (1973)
[27]J R Rabinowitz, T J Swissler, and R Rein, Int J Quantum Chem 86, 353 (1972)
[28]E Clementi and J Mehl, IBMOL -5 program 'Quantum Mechanical Concepts and Algorithms," June 22 1971, IBM Report RJ883
[29]B J Ransil, J Chem Phys 34, 2109 (1961)
[30]N R Kestner, J Chem Phys 48, 252 (1968)
[31]H Margenau and N R Kestner, *Theory of Intermolecular Forces* (Pergamon, Oxford, 1969) Appendix A
[32]H F Schaefer III, D R McLaughlin, F E Harris, and B J Alder, Phys Rev Lett 25, 988 (1970)
[33]P Bertoncini and A C Wahl, Phys Rev Lett 25, 991 (1970)
[34]P E S Wormer, A van Berkel, and A van der Avoird, Mol Phys (to be published)
[35]C A Coulson and P L Davies, Trans Faraday Soc 48, 777 (1952)
[36]E F Haugh and J O Hirschfelder J Chem Phys 23, 1778 (1955)
[37]A Schweig, Int J Quantum Chem 3, 823 (1969)
[38]M Hashimoto, M Hashimoto and T Isobe, Bull Chem Soc Jpn 44, 3230 (1971)
[39]G R Elliot and G E Leroi, J Chem Phys 59, 1217 (1973)
[40]M Hashimoto and T Isobe, Bull Chem Soc Jpn 46, 2581 (1973)
[41]M Hashimoto and T Isobe, Bull Chem Soc Jpn 47, 40 (1974)
[42]C Brecher and R S Halford, J Chem Phys 35, 1109 (1961)
[43]D A Dows, J Chem Phys 36, 2836 (1962)
[44]G Taddei and E Giglio, J Chem Phys 53, 2768 (1970)
[45]E A Matsen, Adv Quantum Chem 1, 60 (1964)
[46]P E S Wormer and A van der Avoird, Int J Quantum Chem 8, 715 (1974)
[47]R McWeeny and B T Sutcliffe *Methods of Molecular Quantum Mechanics* (Academic, London, 1969), p 136
[48]H Margenau and G M Murphy, *The Mathematics of Physics and Chemistry* (Van Nostrand, Princeton, 1956), 2nd ed
[49]M E Rose, J Math and Phys 37, 215 (1958)
[50]A D Buckingham, Q Rev (Lond) 13, 183 (1959)
[51]L Pauling, J Chem Phys 1, 280 (1933)
[52]R McWeeny, Proc R Soc Lond A 223, 306 (1954)
[53]C M Reeves, Thesis, University of Cambridge, England, 1957
[54]I L Cooper and R McWeeny, J Chem Phys 45, 226 (1966)
[55]B T Sutcliffe, J Chem Phys 45, 235 (1966)
[56]F A Matsen A A Cantu, and R D Poshusta, J Phys Chem 70, 1558 (1966)
[57]G F Tantardini M Raimondi, and M Simonetta, Int J Quantum Chem 7, 893 (1973)
[58]W Moffitt, Proc R Soc Lond A 218, 486 (1953)
[59]P O Löwdin Phys Rev 97, 1474 (1955)
[60]R Prosser and S Hagstrom, Int J Quantum Chem 2, 89 (1968)
[61]G A Gallup Int J Quantum Chem 6, 899 (1972)
[62]I G Kaplan and O B Radimova, Int J Quantum Chem 7, 1201 (1973)
[63]C M Reeves, Commun ACM (Assoc Comput Mach) 9, 276 (1966)
[64]J C Slater, J Chem Phys 19, 220 (1951)
[65]We call those orbitals closed-shell orbitals that are doubly occupied in *all* VB structures entering the calculation Orbitals that are singly occupied in one or more structures are called open-shell orbitals
[66]P E S Wormer and A van der Avoird, J Chem Phys 57, 2498 (1972)
[67]This method of dealing with the nonorthogonality was first used by Moffitt[58] and later by G G Balint-Kurti and M Karplus, J Chem Phys 50, 478 (1969)
[68]P O Löwdin, Adv Quantum Chem 5, 185 (1970).
[69]B C Carlson and J M Keller, Phys Rev 105, 102 (1957)
[70]The authors thank Dr E Clementi and Dr F B van Duijneveldt for making available the IBMOL -5 program
[71]E Clementi, Proc Natl Acad Sci USA 60, 2942 (1972), we thank Dr E Clementi and Dr M van Hemert for making available the IBMOL -5A program
[72]A D McLean, "Proceedings of the Conference on Potential Energy Surfaces in Chemistry," Publ RA18, IBM Research, San José, CA, 1971
[73]The authors thank Dr M van Hemert at Leiden University for his four-index transformation program
[74]M Yoshimine, 'The Use of Direct Access Devices in Problems Requiring Reordering of Long Data Lists," Publ RJ555, IBM Research, San José, CA, 1969
[75]R C Raffenetti, Chem Phys Lett 20, 335 (1973)
[76]A Pipano and I Shavitt, unpublished notes, Haifa, June 1965
[77]M Yoshimine, J Comput Phys 11, 449 (1973)
[78]Recently, a description of a computer program written by G H F Diercksen and B T Sutcliffe [Theor Chim Acta 34, 105 (1974)] appeared, which incorporates Reeves' algorithm[63] with a sorting procedure similar to that of Yoshimine [74]
[79]T J Dekker ALGOL 60 Procedures in Numerical Algebra," Mathematical Centre Tracts Amsterdam 1968, Vol 2
[80]R A Faulkner, *Computational Methods in Band Theory*, edited by P M Marcus J F Janak, and A R Williams (Plenum, New York 1971)
[81]S Huzinaga, J Chem Phys 42, 1293 (1965).
[82]S Huzinaga 'Approximate Atomic Functions I," Report of Division of Theoretical Chemistry, The University of Alberta, 1971
[83]T H Dunning Jr, J Chem Phys 55, 716 (1971)
[84]W J Hehre R Ditchfield, and J A Pople J Chem Phys 53, 932 (1970)
[85]H C Allen and E K Plyler, J Am Chem Soc 80, 2673 (1958)
[86]J M Schulman, J W Moskowitz and C Hollister, J Chem Phys 46, 2759 (1967)
[87]I Fischer-Hjalmars and P Siegbahn, Theor Chim Acta 31, 1 (1973)
[88]L C Snyder and H Basch *Molecular Wave Functions and Properties* (Wiley, New York 1972)
[89]A D Buckingham R L Disch, and D A Dunmer, J Am Chem Soc 90, 3104 (1968)
[90]A D King, J Chem Phys 51, 1262 (1969)
[91]It is often found stated in the literature (e g , Ref 31) that

three-body long-range interactions only start appearing in third order of perturbation theory, while the corresponding pair forces occur in second order. This statement is only strictly true for the interaction between neutral atoms in an S state, since its proof as given in Ref. 31, Chap. 5, requires the vanishing of $\langle \hat{\psi}_0^A \hat{\psi}_0^B | V^{AB} | \hat{\psi}_0^A \hat{\psi}_0^B \rangle$. Still, the three-body forces are expected to be much smaller than the pair forces.

[92]R. Polak and J. Paldus, Theor. Chim. Acta **4**, 37 (1966).

[93]O. Nagai and T. Nakamura, Prog. Theor. Phys. **24**, 432 (1960).

[94]T. Kihara, Acta Crystallogr. A **26**, 315 (1970).

[95]J. Felsteiner and D. B. Litvin, Phys. Rev. B **4**, 671 (1971).

[96]1 hartree = 27.21 eV = 627.49 kcal/mol = 2.625×10^6 J/mol, 1 bohr = 0.529167×10^{-10} m. Quadrupole units 1 a.u. = 1.344×10^{-26} esu·cm^2.

[97](CNDO) Molecular Orbital Program by P. A. Dobosh, distributed by Q. C. P. E., Indiana University, Bloomington, IN.

[98]J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory* (McGraw-Hill, New York, 1970).

[99]L. Pauling, *The Nature of the Chemical Bond* (Cornell U. P., Ithaca, NY, 1960).

[100]C. F. Wilcox, S. Winstein, and W. G. McMillan, J. Am. Chem. Soc. **82**, 5450 (1960).

[101]L. W. Flynn and G. Thodos, Am. Inst. Chem. Eng. J. **8**, 362 (1962).

[102]J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954).

[103]H. Basch, Chem. Phys. Lett. **19**, 323 (1973).

[104]A. Johansson, P. Kollman, and S. Rothenberg, Theor. Chim. Acta **26**, 97 (1972).

INTERACTION POTENTIAL FOR He H_2 IN THE REGION OF THE VAN DER WAALS MINIMUM

P J M GEURTS, P E S WORMER and A VAN DER AVOIRD
Institute of Theoretical Chemistry, University of Nijmegen, Nijmegen, The Netherlands

Received 4 July 1975

Interaction energies for the He H_2 system have been computed by a VB method for the intermolecular distances 5.2 $\leq R \leq$ 20.0 bohr and two different orientations of the H_2 molecule (θ = 0° 90°). The results, especially for the anisotropy, are in good agreement with experiment. The dispersion constants C_6 and C_8 and the corresponding anisotropy constants a_6 and a_8 are presented.

1 Introduction

As one of the simplest systems with a multi-dimensional potential surface, the He H_2 complex is of experimental as well as theoretical interest. Integral total collision cross sections [1], spin–lattice relaxation times [2] and Raman line shapes [3] have yielded experimental information on the He H_2 potential surface, and not long ago an analytic intermolecular potential has been proposed by Shafer and Gordon [4] (from here on referred to as SG) which accurately fits all these experimental data.

Theoretical investigations have concentrated either on the short range (repulsive) potential [5–7] or on the long range dispersion forces [8]. Tsapline and Kutzelnigg [9] (henceforth referred to as TK) were the first to calculate van der Waals minima for He H_2. Their minima are close to the experimental results. However, TK predict the linear configuration of the He H_2 system to be favoured by 2.35×10^{-5} hartree above the perpendicular geometry, whereas the empirical results of SG show the linear geometry to be slightly less stable than the perpendicular one (by an amount of 0.3×10^{-5} hartree). Also the theoretical and experimental van der Waals radii are not in complete agreement.

It could be supposed that this discrepancy is caused by the neglect of intra-monomer correlation in the calculations and more specifically by the coupling of intra- with inter-monomer correlation [9, 10]. However, as it seems hard to accept that this correlation effect could be responsible for such a relatively large error in the anisotropy, we thought it worthwhile to perform some of the calculations as well, using another formalism [11–13] and another basis of atomic orbitals, but also neglecting the effects of intra monomer correlation. Interaction curves for two different geometries are presented:

(i) the perpendicular geometry with an angle $\theta = 90°$ between the molecular axis of H_2 and the vector $\boldsymbol{R}$ connecting the midpoint of H_2 with He and

(ii) the linear geometry with $\theta = 0°$.

The distance R has been varied from 5.2 to 20.0 bohr, the H–H distance has been kept constant (1.40 bohr).

2 Method

The method employed in this work is essentially a multistructure valence-bond method based on VB structures containing AO's on He and MO's on H_2. In this formalism it is possible, by using local symmetry, to separate the different terms in the multipole expansion of the dispersion energy [11], although the complete unexpanded form of the interaction operator is used in our calculations. To obtain the contribution from a certain component of the multipole operator the basis must include at least one VB structure representing a locally excited state that combines with the monomer ground state under this multipole component.

Table 1
Atomic orbital basis

	N_{contr} [a]	ζ [b]	Optimization method	Ref.
He s	6		SCF calculation He	[17]
p	2	1.30	van der Waals int. in He_2	[12]
d	1	1.45 [c]	van der Waals int. in He_2	[13]
H s	6	1.2	SCF calculation in H_2 [d]	this work
p	2	2.0	SCF calculation in H_2	this work
p'	2	1.1	van der Waals int. in He H_2	this work
d	1	1.1 [c]	van der Waals int. in He H_2	this work

[a] Number of primitive GTO's in contracted set. [b] Contracted set represents STO with exponent ζ (bohr^{-1}).
[c] Exponent GTO $\alpha = 0.2738$ (bohr^{-2}). [d] Contraction coefficients optimized. [e] Exponent GTO $\alpha = 0.1576$ (bohr^{-2}).

Within the VB framework one can define two quantities $\Delta I^{(1)}$ and $\Delta I^{(2)}$ resembling first- and second-order perturbation energies, respectively. $\Delta I^{(1)}$ is the expectation value of the total hamiltonian over the dimer ground state minus the SCF energies of the free monomers. $\Delta I^{(2)}$ is the difference between the VB multi-structure interaction energy and the just defined first-order energy. This difference is asymptotically equal to London's well-known second order energy expression [11].

The orbitals on the different monomers do of course overlap, but since we have found previously that a well-chosen orthogonalization does not influence the results to a significant extent, we have orthogonalized the basis, while meeting the following requirements:

the ground state VB structure, and with it $\Delta I^{(1)}$, is left invariant,

the orbitals remain as localized as possible on the respective monomers, consequently they reflect the symmetry of the subsystems.

A careful orthogonalization is of the utmost importance. If, for instance, one allows all orbitals in the dual space to mix among each other, as is done in the method of bi-orthogonal orbitals [14], one finds first-order energies which are very sensitive to the basis set and which sometimes are negative in regions of the potential surface where strong repulsions are expected [14].

3. Basis

The selection of an orbital basis needs special care in the study of van der Waals interactions, because these are usually very small. It has been found [13], for instance, that the orbitals constituting the excited states must be optimized by maximizing the dispersion energy. Good ground state orbitals are also required for a reliable estimate of dispersion energies [13] as well as for correct exchange repulsions [15]. Thus, we have included in the AO basis two p-orbitals on each hydrogen, one rather diffuse orbital to obtain the corresponding parts of the dispersion energy and another more compact orbital necessary for the description of the exchange repulsion between He and H_2. The need for at least two p-orbitals in cases such as this has sometimes been overlooked [16]. Contrary to the case of the p-orbitals, one d orbital per hydrogen atom appears to be sufficient, mainly because the orbital exponent required for getting a good ground state energy of H_2 very nearly optimizes the dispersion energy of He H_2. One must realize, in this connection, that the dispersion energy is not very sensitive to variations of this exponent in the neighbourhood of the optimum.

The inclusion of an optimized p-orbital on He is necessary to account for the dipole excitations on this atom; one optimized d-orbital on He takes care of the quadrupole excitations.

The AO's used in this work are contracted GTO's with tesseral harmonics as their angular parts. Except for the 1s-orbitals on H and He, they have been fitted to STO's with exponents ζ. The ζ-values have subsequently been optimized. See table 1 for a summary of the basis. We have also experimented with larger contractions of the He and H p-orbitals, but this hardly affected the results.

The MO's of σ symmetry on H_2 are obtained from

Table 2

	Present work	TK [9]	HF limit
SCF energy He [a]	2.8611163	2.8614912	2.86168 [c]
SCF energy H_2 [a]	1.1330237	1.1329092	-1.133629 [b]
quadrupole H_2 [d]	0.4931		0.4933 [b]

[a] Hartree. [b] Ref. [18]. [c] Ref. [17].
[d] $\frac{1}{2}(3\langle z^2\rangle - \langle r^2\rangle)$ in au

Table 3
First-order, second-order and total interaction energies. Distance in bohr, energy in 10^{-5} hartree

	Linear geometry, $\theta = 0°$			Perpendicular geometry, $\theta = 90°$		
R	$\Delta E^{(1)}$	$\Delta E^{(2)}$	ΔE_{tot}	$\Delta E^{(1)}$	$\Delta E^{(2)}$	ΔE_{tot}
5.2	56.50	28.32	28.18	38.87	-22.13	16.74
5.6	24.40	18.71	5.69	16.79	14.91	1.88
6.0	10.41	12.59	-2.18	7.15	10.16	-3.01
6.3	5.45	9.49	-4.04	3.73	7.68	-3.95
6.5	3.52	-7.91	4.39	2.41	6.41	-4.00
6.6	2.83	-7.23	4.40	1.93	5.85	-3.92
7.0	1.16	-5.10	3.94	0.78	-4.11	3.33
8.0	0.11	2.26	2.15	0.07	-1.81	-1.74
9.0	0.00	1.08	1.08	0.00	-0.87	0.87
10.0	0.00	0.56	0.56	0.00	0.46	-0.46
11.0	0.00	-0.32	0.32	0.00	-0.26	0.26
12.0	0.00	0.18	-0.18	0.00	-0.15	0.15
20.0	0.00	0.01	-0.01	0.00	0.02	0.02

an SCF calculation on the free molecule, and those of π- and δ-symmetry are simply symmetric and antisymmetric combinations of equivalent orbitals on the atoms.

As in earlier work [11-13] we have found again that VB structures representing the coupling of triplet excited monomers can be omitted. This halves the number of VB structures contributing to the dispersion energy. In total, about 50 of such singlet–singlet structures can be derived from the given orbital basis, all of these have been taken into account in this work.

The quality of the atomic orbital basis may be judged from the values in table 2. Another criterion for the adequacy of the basis is the "saturation test" [19], which is a computation of the SCF energy of each of the single subsystems in the dimer basis. We have found the following small energy improvements at $R = 6.5$ bohr, expressed in 10^{-5} hartree 0.34 and

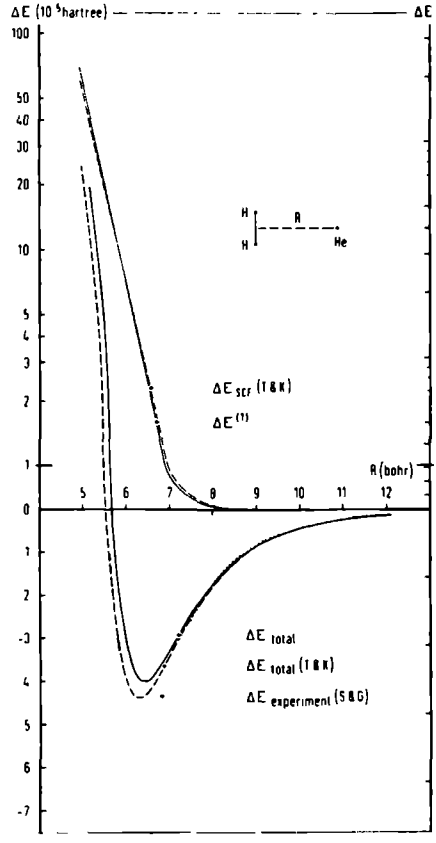


Fig. 1. Perpendicular geometry. Total interaction energy ΔE_{tot} of TK [9], SG [4] and this work. First-order energy $\Delta E^{(1)}$ of this work, SCF-interaction energy ΔE_{SCF} of TK [9]. From 1×10^{-5} hartree upwards the energy scale is logarithmic.

0.26 for He, 0.16 and 0.08 for H_2 in the case of the linear and the perpendicular geometry, respectively. If we had obtained the first-order interaction from SCF calculations on the dimer, we would have had to correct for these small effects, but because we calculate $\Delta E^{(1)}$ directly from the monomer orbitals, the fact that the basis set is not completely saturated do not concern us.

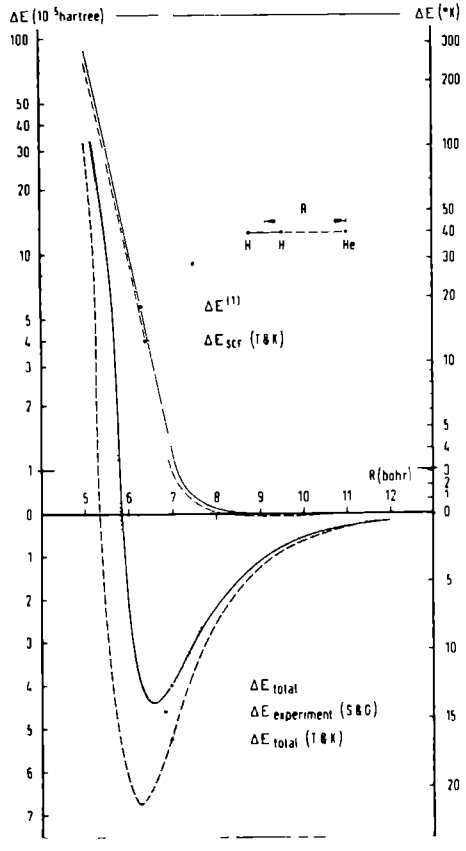


Fig. 2 Linear geometry. Total interaction energy ΔE_{tot} of TK [9], SG [4] and this work. First-order energy $\Delta E^{(1)}$ of this work. SCF-interaction energy ΔE_{SCF} of TK [9]. From 1×10^{-5} hartree upwards the energy scale is logarithmic

4 Results and discussions

In table 3 the first-order, second-order and total interaction energies are given for the two configurations considered in this work. Figs. 1 and 2 give some of these results graphically, showing also a comparison with the total interaction energies of SG [4] and the SCF and total interaction energies of TK [9].

4.1 First order energy

For the perpendicular geometry one notes that our first-order energies are very similar to the SCF energies of TK, the hardly significant difference can probably be explained by the differences in the basis sets. In the case of the linear geometry, our first order results differ a little more from the SCF curve of TK, which is lying lower. The induction part (mainly permanent quadrupole on H_2, induced dipole on He) of the interaction energy could be a source of this discrepancy, because it is contained in the dimer SCF energy but not in our first order energy. We have therefore computed the induction energy explicitly, by the VB method, as well as classically using the quadrupole moment of H_2 from table 2 and the polarizability $\alpha = 1.393$ (bohr)^3 of He [13]. We find that the two methods agree in all significant figures, which are not many in this case because the effect is very small indeed: 0.05×10^{-5} hartree at 6.5 bohr (linear geometry). In any event this value is too small to explain the difference between TK's and our results. Possibly TK have improved the SCF energies of the subsystems somewhat in the dimer basis, due to a non-saturated monomer basis and have not subsequently corrected for this effect. This allegation is supported by the occurrence of a minimum in their curve which, although very weak, is still too deep to be caused by induction. Also our first-order energy at 5.2 bohr compares quite well with the best SCF value of ref. [5] after correction for the induction energy: 56.5×10^{-5} hartree versus 56.3×10^{-5} hartree [5], whereas TK find about 51×10^{-5} hartree (interpolated) for the same value. (For the perpendicular case with the same distance these values are 38.9 (this work), 39.6 [5] and about 37 (TK).)

The first order repulsion can be fitted quite satisfactorily by the following potential valid for $R \geq 5.2$ bohr. (The θ dependency is given by a second order Legendre polynomial.)

$$I^{(1)}(R,\theta) = A e^{-\beta R}[1 + \gamma P_2(\cos\theta)]$$

The optimal parameters are

$$A = 34.5 \text{ hartree}, \quad \beta = 2.16 \text{ (bohr)}^{-1}, \quad \gamma = 0.267$$

The β value is somewhat larger than any of the values proposed in ref. [5] for $3.8 \leq R \leq 5.2$ (bohr), a range which is different from ours though. Since a

Table 4

	Linear geometry			Perpendicular geometry		
	this work	TK [9]	SG [4]	this work	TK	SG
R_m (bohr)	6.58	6.25	6.58	6.42	6.29	6.30
E_m (10^{-5} hartree)	4.40	6.75	4.77	4.02	4.40	5.06

small deviation in β gives a considerable effect on A, it is not surprising that our A-value differs by a factor of about 3 from the values of ref. [5].

4.2. Total energy

With regard to the total energy, we see from table 4 that for the perpendicular geometry TK's calculations agree better with the experimental results than ours, although our values do not deviate too much either. We have observed before that our method has a tendency to underestimate the dispersion energy to some extent. Using a comparable basis we have found for He–He [13] a dispersion energy $\Delta E^{(2)}$ which is 90% of the value computed in a much larger basis including f-orbitals [20].

As to the linear geometry, the results of this work are in full agreement with the experimental R_m-value and, again, underestimate the experimental well depth somewhat, whereas TK find too deep a minimum at too short a distance (table 4)*.

So, although the anisotropy found in this work is much less pronounced than that of TK [9], we still predict the wrong geometry to be more stable, that is, comparing with SG [4]. To explain this discrepancy one may point out several inaccuracies in this work, such as the neglect of intra-monomer correlation or a possible geometry-dependent underestimate of the dispersion, but it must also be noted that SG's potential lacks some flexibility in the long range part. The following discussion may clarify this remark. Within the VB formalism one can calculate the dispersion coefficients C_6, C_8, etc., directly. Thus computing C_6 and C_8 for the linear and perpendicular geometry, one may extract averaged C_6- and C_8-values plus the corresponding anisotropy constants α_6 and α_8. In this manner we have computed the following long range potential:

$$V^{(2)}(R,\theta) = -C_6 R^{-6} [1 + \alpha_6 P_2(\cos\theta)] - C_8 R^{-8} [1 + \alpha_8 P_2(\cos\theta)] ,$$

with

$$C_6 = 4.34 \text{ au}, \quad C_8 = 49 \text{ au}, \quad \alpha_6 = 0.149, \quad \alpha_8 = 0.23 .$$

The values of C_6, C_8 and α_6 are in fair agreement with the values $C_6 = 4.01$ au, $C_8 = 41$ au, $\alpha_6 = 0.105$ quoted by SG. This potential gives an excellent fit to the second-order energies of table 3 for $R \geq 7.0$ bohr. Now, SG assume α_6 and α_8 to be identical, thus overestimating C_8 for the perpendicular geometry, while underestimating C_8 for the linear case. So, by this lack of flexibility in their long range anisotropy they favour the perpendicular geometry somewhat above the linear one, but because of the way they fit the potential, it is difficult to see how this affects the minima.

It could be presumed that the superposition, $V^{(1)} + V^{(2)}$, of the short and the long range potential can yield a reasonable fit for the whole range. This is not so, mainly because of the inadequacy of $V^{(2)}$ to represent charge-penetration effects. A better fit would require some extra (exponential) terms to account for charge penetration.

Since scattering experiments are often interpreted by the use of Lennard-Jones potentials [21], we have computed the following analytic form from the positions and the depths of the minima

$$V_{L\text{-}J}(R,\theta) = -2\epsilon(R_0/R)^6 [1 + q_{2,6} P_2(\cos\theta)] + c(R_0/R)^{12} [1 + q_{2,12} P_2(\cos\theta)] ,$$

* When informed about our results Professor Kutzelnigg has communicated that their potential curves become very similar to ours when they extended their formalism in order to account for the coupling between inter- and intra-monomer correlations. In comparing these results it must be remembered though, that the relative contributions from inter- and intra-monomer terms depend on the degree of localization of the orbitals on the subsystems.

with

$\epsilon = 4.13 \times 10^{-5}$ hartree, $R_0 = 6.48$ bohr,

$q_{2,6} = 0.164$, $q_{2,12} = 0.271$.

The L-J potential gives a good description of the computed van der Waals well, but overestimates both the repulsion and the long range attraction in the region considered. The anisotropy constants $q_{2,12}$ and $q_{2,6}$ are not very different from our corresponding asymptotic values $\gamma = 0.267$ and $\alpha_6 = 0.149$. Furthermore, they are in good agreement with experimental results [22] obtained from the scattering of molecular H_2 beams with several of the noble gases.

Acknowledgement

We thank Dr. J. Reuss for suggesting the problem, and him and Drs. L. Zandee, T. van Berkel and F. Mulder for valuable discussions.

References

[1] R. Gengenbach and Ch. Hahn, Chem. Phys. Letters 15 (1972) 604.
[2] J.W. Riehl, C.J. Fisher, J.D. Baloga and J.L. Kinsey, J. Chem. Phys 58 (1973) 4571.
[3] B.K. Gupta and A.D. May, Can. J. Phys 50 (1972) 1747.
[4] R. Shafer and R.G. Gordon, J. Chem. Phys. 58 (1973) 5422.
[5] M.D. Gordon and D. Secrest, J. Chem. Phys. 52 (1970) 120
[6] M. Krauss and F.M. Mies, J. Chem. Phys. 42 (1965) 2703.
[7] C.S. Roberts, Phys. Rev. 131 (1963) 203
[8] P. Langhoff, R.G. Gordon and M Karplus, J. Chem. Phys. 55 (1971) 2126, and references therein.
[9] B. Tsapline and W. Kutzelnigg, Chem. Phys. Letters 23 (1973) 173.
[10] B Liu and A.D. McLean, J. Chem. Phys. 59 (1973) 4557.
[11] P.L.S. Wormer and A. van der Avoird, J Chem. Phys. 62 (1975) 3326.
[12] P.E.S. Wormer, T. van Berkel and A. van der Avoird, Mol. Phys. 29 (1975) 1181.
[13] F. Mulder, P. Geurts and A. van der Avoird, Chem. Phys. Letters 33 (1975)215.
[14] F. Kochanski and J.F. Gouyet, Mol. Phys. 29 (1975) 693.
[15] J.G.C M. van Duijneveldt-van de Rijdt and F.B. van Duijneveldt, Chem. Phys. Letters 17 (1972) 425
[16] F. Kochanski, J. Chem. Phys. 58 (1973) 5823.
[17] S. Huzinaga, J Chem. Phys. 42 (1965) 1293.
[18] W. Kołos and C.C J. Roothaan, Rev. Mod. Phys. 32 (1960) 219.
[19] W. Kutzelnigg, V Staemmler and C. Hoheisel, Chem. Phys. 1 (1973) 27.
[20] D.R. McLaughlin and H.F. Schaefer III, Chem. Phys. Letters 12 (1971) 244.
[21] H. Moerkerken, L. Zandee and J. Reuss, Chem. Phys. 11 (1975), to be published.
[22] L. Zandee and J. Reuss, private communication.

SUMMARY

The most important part of this dissertation consists of reprints of papers published earlier. Two subjects are treated in these papers:

I. Application of the representation theory of GL(n) and S_N to the many-body problem.

II. Calculations on intermolecular forces.

Many of the techniques discussed under I are applied in II.

I. In the chapters I.1. through I.6. of this dissertation the close connection is discussed which exists between spin, permutation symmetry and transformation properties of N-electron wave functions. This mathematical connection is treated in detail, mainly because this theory is not yet very well-known among chemists. In two of the reprinted papers the results of the theory are applied to the transformation properties of N-electron wave functions with respect to "blocked" orbital mixings.

The other two group theoretically oriented papers treat respectively the group theoretical classification of atomic Russell-Saunders states and the classification of more-particle interactions with respect to permutation symmetry.

II. The goal of these investigations has been the development of a practical method for the computation of intermolecular potentials in the region of the Van der Waals minimum. The first article on this subject presents the results of calculations by means of a simple semi-empirical method (the CNDO method). The results were such that we looked for a more reliable method. This was found in the multi-structure valence bond formalism. The first paper based on this formalism (calculations on He_2) presents a test on the reliability and the tractability of the method. The second article is an attempt to furnish more insight into the bonding of a chemically more interesting complex (the ethylene dimer). In the third valence bond paper we have attempted to calculate a potential of good quality for $He-He_2$. An important reason for the choice of the $He-H_2$ complex was the fact that experimental work on this complex was under way in the atomic and molecular spectroscopy group of this university. So the computed potential could be compared directly with the experiment.

Because the VB calculations were guided by London's theory of intermolecular forces, the papers are preceded by a presentation of this theory in terms of irreducible tensor operators.

SAMENVATTING

Dit proefschrift bestaat voor het voornaamste gedeelte uit herdrukken van reeds eerder verschenen artikelen. Deze artikelen betreffen twee onderwerpen:

I. Toepassing van de representatietheorie van GL(n) en S_N op het veeldeeltjes probleem.

II. Berekeningen aan intermoleculaire wisselwerkingen.

Vele van de technieken besproken onder I zijn toegepast in II.

I. Zoals uiteengezet wordt in de hoofdstukken I.1 t/m I.6 van deze dissertatie, bestaat er een nauw verband tussen spin, permutatiesymmetrie en transformatie-eigenschappen van N-elektron golffuncties. Dit wiskundig verband wordt in detail behandeld, voornamelijk omdat deze theorie voor een groot gedeelte nog nauwelijks bekend is onder chemici. In twee van de herdrukte artikelen worden de resultaten van de theorie toegepast op de transformatie-eigenschappen van N-elektron golffuncties onder "geblokte" orbital mengingen.

De andere twee groepentheoretisch georiënteerde artikelen betreffen respectievelijk de groepentheoretische klassificatie van atomaire Russell-Saunders toestanden en de klassificatie van meer-deeltjes interacties met betrekking tot permutatiesymmetrie.

II. Doel van dit onderzoek was te komen tot een bruikbare methode voor de berekening van intermoleculaire potentialen in het gebied van het Van der Waals minimum. Het eerste artikel over dit onderwerp geeft de resultaten van berekeningen met een eenvoudige semi-empirische methode (de CNDO methode). Deze resultaten waren zodanig dat gezocht werd naar een meer betrouwbare methode. Deze werd gevonden in het "multi-structure valence bond" formalisme. Het eerste artikel dat gebaseerd is op dit formalisme (berekeningen aan He_2) presenteert een test op de betrouwbaarheid en handelbaarheid van de methode. Het tweede artikel is een poging meer inzicht te verschaffen in de binding van een chemisch interessanter complex (het etheen dimeer). In het derde valence bond artikel is geprobeerd een potentiaal voor $He-H_2$ van goede kwaliteit te berekenen. Een voorname reden voor de keuze van het complex $He-H_2$ was dat hieraan experimenteel werk werd verricht op de afdeling atoom- en molecuulspectroscopie van deze universiteit, en dat de berekende potentiaal dus

direct aan het experiment getoetst kon worden.

Omdat de VB berekeningen London's theorie voor intermoleculaire krachten als leidraad hadden, worden de artikelen voorafgegaan door een presentatie van deze theorie in termen van niet-reduceerbare tensoroperatoren.

ACKNOWLEDGEMENTS

I thank Dr. M. Boon for many enlightning discussions about the theory described in the first part of this thesis and Prof. F.A. Matsen for his stimulating lectures on the same subject.

Many people have been of great help to the calculations reported on in the second part of this thesis. Among those not already properly thanked in the acknowledgements of the reprinted papers, I must especially mention Prof. D.B. Chesnut, who made my stay at Duke University pleasant and worthwhile, the staff of the computing center in Nijmegen, particularly Chris Meijer, who taught me how to operate the computer and was always ready to help when problems arose, and many students, especially Ton van Berkel and Petro Geurts, who collaborated on two of the papers reprinted in this work.

Finally I thank Jenny Wormer for the enormous task of typing all the parts of this thesis that are not in the form of reprinted articles, and Ine van Berkel and the offset department for typing the manuscripts of the papers reprinted in this work.

The investigations were supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

CURRICULUM

De schrijver van dit proefschrift bezocht de Technische Hogeschool te Delft, Duke University te Durham N.C., V.S. (gedurende een jaar), en is sinds 15 september 1970 werkzaam aan het Instituut voor Theoretische Chemie van de Universiteit van Nijmegen.

STELLINGEN

I

Een van de voornaamste uitkomsten van ETBM berekeningen aan de kernbewegingen in het He-trimeer, nl. het resultaat dat de gemiddeld kernafstand 23 Å is, moet op fysische en mathematische gronden verworpen worden.

T.K. Lim, M.A. Zuniga, J. Chem. Phys. 63, 2245 (1975).

II

Auteurs die het chemisorptiegedrag van overgangsmetalen proberen te voorspellen aan de hand van de eigenschappen van het ongestoorde oppervlak, via het "surface molecule" model, doen vaak onvoldoende recht aan dit model, omdat ze geen rekening houden met de sterke chemische binding die kan optreden bij adsorptie.

D.R. Penn, Surface Sci. 39, 333 (1975).
M.J. Kelly, Surface Sci. 43, 587 (1974);
J. Phys. C7, L157 (1974).

III

De metingen van Kessener en Reuss aan de anisotropie van de werkzame doorsnede van het systeem NO-N_2 zouden een aanwijzing kunnen zijn dat dipool-quadrupool- en quadrupool-quadrupool-interacties de glory interferenties volledig "quenchen". Deze interacties kunnen namelijk het Van der Waals minimum in bepaalde oriëntaties doen verdwijnen.

H.P.M. Kessener, J. Reuss, Chem. Phys. Lett. 31, 212 (1975).
Dit proefschrift.

IV

Het bewijs van Margenau en Kestner dat drie-deeltjes-interacties pas beschreven kunnen worden met derde en hogere orde storingsrekening is slechts geldig voor atomen in een S-toestand.

H. Margenau, N.R. Kestner, Theory of Intermolecular Forces,
Pergamon Press, Oxford 1969.

V

Het is verbazingwekkend dat Xα berekeningen aan het Ar-kristal een roosterenergie en een roosterafstand vrijwel gelijk aan de experimentele waarden geven, terwijl uit berekeningen aan het Ne-dimeer geconcludeerd wordt dat de Xα-methode niet in staat is Van der Waals interacties te beschrijven.

D.D. Konowalow, P. Weinberger, J.L. Calais, J.W.D. Conolly, Chem. Phys. Lett. 16, 81 (1972).
S.B. Trickey, F.W. Averill, F.R. Green jr., Phys. Lett. 41A, 385 (1972).

VI

Het toch al naïeve CFSO-BEBO model voor de beschrijving van chemisorptie op overgangsmetalen neemt ten onrechte aan dat de som van de bondorders in Pt_2-X_2 complexen altijd gelijk twee is, ongeacht de bindingsafstanden.

W.H. Weinberg, R.P. Merrill, Surface Sci. 33, 493 (1972).
W.H. Weinberg, H.A. Deans, R.P. Merrill, Surface Sci. 41, 312 (1974).

VII

Een van de conclusies van het advies van de gezondsheidsraad luidt dat de biologisch effectieve stralingsdosis welke de Nederlandse bevolking in het jaar 2000 zal ontvangen (na algemene invoering van kernenergie), niet noemenswaard groter zal zijn dan momenteel. Deze conclusie is aanvechtbaar omdat de totale hoeveelheid radioactieve isotopen in de biosfeer wél zeer aanzienlijk zal toenemen. Daarom zal eerst meer onderzoek verricht moeten worden naar het effect van straling op gehele eco-systemen voordat men besluit tot toepassing van kernenergie op grote schaal.

Advies Gezondheidsraad: "Kerncentrales en Volksgezondheid; invloed van de kernenergie op volksgezondheid en milieu in Nederland bij een totale capaciteit van 3500 MWe". September 1975.

Nijmegen 11 december 1975 P.E.S. Wormer